



(19)

Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 880 590 B1

(12)

EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention
of the grant of the patent:
29.10.2004, Bulletin 2004/43

(51) Int Cl.7: **C12N 15/55**, C12N 9/16,
C12N 1/21, C12P 13/02,
C12Q 1/68, C07K 16/40,
G01N 33/53

(21) Application number: 97906528.1

(22) Date of filing: 11.02.1997

(86) International application number:
PCT/US1997/002039

(87) International publication number:
WO 1997/030160 (21.08.1997 Gazette 1997/36)

(54) ESTERASES

ESTERASEN

ESTERASES

(84) Designated Contracting States:
**AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC
NL PT SE**

(30) Priority: 16.02.1996 US 602359

(43) Date of publication of application:
02.12.1998 Bulletin 1998/49

(60) Divisional application:
04017203.3

(73) Proprietor: DIVERSA CORPORATION
San Diego, California 92121 (US)

(72) Inventors:

- ROBERTSON, Dan, E.
Haddonfield, NJ 08033 (US)
 - MURPHY, Dennis
Paoli, PA 19301 (US)
 - REID, John
Bryn Mawr, PA 19010 (US)
 - MAFFIA, Anthony, M.
Wilmington, DE 19810 (US)
 - LINK, Steven
Wilmington, DE 19803 (US)
 - SWANSON, Ronald, V.
Media, PA 19063 (US)
 - WARREN, Patrick, V.
Philadelphia, PA 19163 (US)
 - KOSMOTKA, Anna
Brookhaven, PA 19015 (US)
 - CALLEN, Walter
San Diego, CA 92117 (US)

(74) Representative: VOSSIUS & PARTNER
Siebertstrasse 4
81675 München (DE)

(56) References cited:
EP-A-0 318 775 WO-A-93/25691
WO-A-94/02617 WO-A-95/23853
WO-A-95/34644 US-A-5 306 636

- WRAY L. ET AL.: "Streptomyces coelicolor transcription factor, AC L03213" EBI DATABASE, 9 October 1992 (1992-10-09), XP002145895
 - KUGIMIYA WATARU ET AL: "Molecular cloning and structure of the gene for esterase from a thermophilic bacterium, Bacillus stearothermophilus IFO 12550." BIOSCIENCE BIOTECHNOLOGY AND BIOCHEMISTRY 1992, vol. 56, no. 12, 1992, pages 2074-2075, XP002138593 ISSN: 0916-8451
 - HUDDLESTON S ET AL: "The identification and partial characterisation of a novel inducible extracellular thermostable esterase from the archaeon Sulfolobus shibatae." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS 1995, vol. 216, no. 2, 1995, pages 495-500, XP002138594 ISSN: 0006-291X

EP 0 880 590 B1

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

- DERWENT BIOTECHNOLOGY ABSTRACTS, Vol. 14, No. 15, 17 July 1995, page 125, Abstract No. 95-09223, LADRAT C. et al., "Thermostable Enzymes Screened on Thermophilic Microorganisms from Deep-Sea Hydrothermal Vents --beta- Glucosidase, Alcohol Dehydrogenase, Protease and Esterase Activity"; & C.R. SEANCES ACAD. SCI. SER 3., 1995, 318(4), 423-429.
- DERWENT BIOTECHNOLOGY ABSTRACTS, Vol. 13, No. 15, 25 July 1994, pages 95-96, Abstract No. 94-08869, CORNEC L. et al., "Thermostable Esterase Activity from Deep-Sea Sulfur Dependent Thermophilic Archaeabacteria-Thermostable Enzyme Purification from Thermophilic Archaeabacterium and Characterization"; & THERMOPHILES SCI. TECHNOL., 1992, 18.
- APPL. MICROBIOL. BIOTECHNOL., 1994, Vol. 42, CANGANELLA F., "Characterization of Amylolytic and Pullulytic Enzymes from Thermophilic Archea and from a New Fervidobacterium Species", pages 239-245.

Description

[0001] This invention relates to newly identified polynucleotides, polypeptides encoded by such polynucleotides, the use of such polynucleotides and polypeptides, as well as the production and isolation of such polynucleotides and polypeptides. More particularly, the polynucleotides and polypeptides of the present invention have been putatively identified as esterases. Esterases are enzymes that catalyze the hydrolysis of ester groups to organic acids and alcohols.

[0002] Many esterases are known and have been discovered in a broad variety of organisms, including bacteria, yeast and higher animals and plants. A principal example of esterases are the lipases, which are used in the hydrolysis of lipids, acidolysis(replacement of an esterified fatty acid with a free fatty acid) reactions, transesterification(exchange of fatty acids between triglycerides)reactions, and in ester synthesis. The major industrial applications for lipases include: the detergent industry, where they are employed to decompose fatty materials in laundry stains into easily removable hydrophilic substances; the food and beverage industry where they are used in the manufacture of cheese, the ripening and flavoring of cheese, as antistaling agents for bakery products, and in the production of margarine and other spreads with natural butter flavors; in waste systems; and in the pharmaceutical industry where they are used as digestive aids.

[0003] The polynucleotides and polypeptides of the present invention have been identified as esterases as a result of their enzymatic activity.

[0004] In accordance with one aspect of the present invention, there are provided novel enzymes, as well as active fragments, analogs and derivatives thereof.

[0005] In accordance with another aspect of the present invention, there are provided isolated nucleic acid molecules encoding the enzymes of the present invention including mRNAs, cDNAs, genomic DNAs as well as active analogs and fragments of such enzymes.

[0006] In accordance with yet a further aspect of the present invention, there is provided a process for producing such polypeptides by recombinant techniques comprising culturing recombinant prokaryotic and/or eukaryotic host cells, containing a nucleic acid sequence of the present invention, under conditions promoting expression of said enzymes and subsequent recovery of said enzymes.

[0007] In accordance with yet a further aspect of the present invention, there is provided a process for utilizing such enzymes, or polynucleotides encoding such enzymes for hydrolyzing ester groups to yield an organic acid and an alcohol. The esterases of the invention are stable at high temperatures and in organic solvents and, thus, are superior for use in production of optically pure chiral compounds used in pharmaceutical, agricultural and other chemical industries.

[0008] In accordance with yet a further aspect of the present invention, there are also provided nucleic acid probes comprising nucleic acid molecules of sufficient length to hybridize to a nucleic acid sequence of the present invention.

[0009] In accordance with yet a further aspect of the present invention, there is provided a process for utilizing such enzymes, or polynucleotides encoding such enzymes, for *in vitro* purposes related to scientific research, for example, to generate probes for identifying similar sequences which might encode similar enzymes from other organisms by using certain regions, *i.e.*, conserved sequence regions, of the nucleotide sequence.

[0010] These and other aspects of the present invention should be apparent to those skilled in the art from the teachings herein.

[0011] The following drawings are illustrative of embodiments of the invention and are not meant to limit the scope of the invention as encompassed by the claims.

[0012] Figure 1 is an illustration of the full-length DNA (SEQ ID NO:23) and corresponding deduced amino acid sequence (SEQ ID NO:33) of *Staphylothermus marinus* F1-12LC of the present invention. Sequencing was performed using a 378 automated DNA sequencer (Applied Biosystems, Inc.) for all sequences of the present invention.

[0013] Figure 2 is an illustration of the full-length DNA (SEQ ID NO:24) and corresponding deduced amino acid sequence (SEQ ID NO:34) of *Pyrodictium* TAG 11-17LC.

[0014] Figure 3 is an illustration of the full-length DNA (SEQ ID NO:25) and corresponding deduced amino acid sequence (SEQ ID NO:35) of *Archaeoglobus venificus* SNP6-24LC.

[0015] Figure 4 is an illustration of the full-length DNA (SEQ ID NO:26) and corresponding deduced amino acid sequence (SEQ ID NO:36) of *Aquifex pyrophilus*-28LC.

[0016] Figure 5 is an illustration of the full-length DNA (SEQ ID NO:27) and corresponding deduced amino acid sequence (SEQ ID NO:37) of M11TL-29L.

[0017] Figure 6 is an illustration of the full-length DNA (SEQ ID NO:28) and corresponding deduced amino acid sequence (SEQ ID NO:38) of *Thermococcus* CL-2-30LC.

[0018] Figure 7 is an illustration of the full-length DNA (SEQ ID NO:29) and corresponding deduced amino acid sequence (SEQ ID NO:39) of *Aquifex* VF5-34LC.

[0019] Figure 8 is an illustration of the full-length DNA (SEQ ID NO:30) and corresponding deduced amino acid

sequence (SEQ ID NO:40) of *Teredinibacter*-42L.

[0020] Figure 9 is an illustration of the full-length DNA (SEQ ID NO:31) and corresponding deduced amino acid sequence (SEQ ID NO:41) of *Archaeoglobus fulgidus* VC16-16MC.

5 [0021] Figure 10 is an illustration of the full-length DNA (SEQ ID NO:32) and corresponding deduced amino acid sequence (SEQ ID NO:42) of *Sulfolobus solfataricus* P1-8LC.

[0022] Figure 11 is an illustration of the full-length DNA (SEQ ID NO:43) and corresponding deduced amino acid sequence (SEQ ID NO:53) of LA11.1 Esterase es2 of the present invention.

10 [0023] Figure 12 is an illustration of the full-length DNA (SEQ ID NO:44) and corresponding deduced amino acid sequence (SEQ ID NO:54) of Whale Mat Sample 11.801 Esterase es9.

[0024] Figure 13 is an illustration of the full-length DNA (SEQ ID NO:45) and corresponding deduced amino acid sequence (SEQ ID NO:55) of *Metallosphaera Prunae Ron* 12/2 Esterase 23mc1.

15 [0025] Figure 14 is an illustration of the full-length DNA (SEQ ID NO:46) and corresponding deduced amino acid sequence (SEQ ID NO:56) of *Thermotoga neapolitana* 5068 Esterase 56mc4.

[0026] Figure 15 is an illustration of the full-length DNA (SEQ ID NO:47) and corresponding deduced amino acid sequence (SEQ ID NO:57) of *Melittangium lichenicola* Esterase 77mc1.

20 [0027] Figure 16 is an illustration of the full-length DNA (SEQ ID NO:48) and corresponding deduced amino acid sequence (SEQ ID NO:58) of Whale Mat Sample 11.801 Esterase es2.

[0028] Figure 17 is an illustration of the full-length DNA (SEQ ID NO:49) and corresponding deduced amino acid sequence (SEQ ID NO:59) of Whale Mat Sample AD3059 Esterase es4.

25 [0029] Figure 18 is an illustration of the full-length DNA (SEQ ID NO:50) and corresponding deduced amino acid sequence (SEQ ID NO:60) of *Microscilla furvescens* Esterase 53sc2.

[0030] Figure 19 is an illustration of the full-length DNA (SEQ ID NO:51) and corresponding deduced amino acid sequence (SEQ ID NO:61) of *Thermotoga maritima* MSB8 Esterase 6sc1.

30 [0031] Figure 20 is an illustration of the full-length DNA (SEQ ID NO:52) and corresponding deduced amino acid sequence (SEQ ID NO:62) of *Polyangium brachysporum* Esterase 78mc1.

[0032] The term "gene" means the segment of DNA involved in producing a polypeptide chain; it includes regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons).

35 [0033] A coding sequence is "operably linked to" another coding sequence when RNA polymerase will transcribe the two coding sequences into a single mRNA, which is then translated into a single polypeptide having amino acids derived from both coding sequences. The coding sequences need not be contiguous to one another so long as the expressed sequences ultimately process to produce the desired protein.

[0034] "Recombinant" enzymes refer to enzymes produced by recombinant DNA techniques; i.e., produced from cells transformed by an exogenous DNA construct encoding the desired enzyme. "Synthetic" enzymes are those prepared by chemical synthesis.

40 [0035] A DNA "coding sequence of" or a "nucleotide sequence encoding" a particular enzyme, is a DNA sequence which is transcribed and translated into an enzyme when placed under the control of appropriate regulatory sequences.

[0036] In accordance with an aspect of the present invention, there are provided isolated nucleic acids (polynucleotides) which encode for the mature enzymes having the deduced amino acid sequences of Figures 1-20 (SEQ ID NOS:23-32 and 43 to 52.)

45 [0037] The deposit(s) have been made under the terms of the Budapest Treaty on the International Recognition of the deposit of micro-organisms for purposes of patent procedure. The strains will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. These deposits are provided merely as convenience to those of skill in the art and are not an admission that a deposit would be required under 35 U.S.C. §112. The sequences of the polynucleotides contained in the deposited materials, as well as the amino acid sequences of the polypeptides encoded thereby, are controlling in the event of any conflict with any description of sequences herein. A license may be required to make, use or sell the deposited materials, and no such license is hereby granted.

50 [0038] The polynucleotides of this invention were originally recovered from genomic gene libraries derived from the following organisms:

[0039] *Staphylothermus marinus* F1 is a thermophilic sulfur archaea which was isolated in Vulcano, Italy. It grows optimally at 85°C ($T_{max} = 98^\circ\text{C}$) at pH 6.5.

[0040] *Pyrodictium TAG11* is a thermophilic sulfur archaea which was isolated in the Middle Atlantic Ridge. It grows optimally at 103°C ($T_{max} = 110^\circ\text{C}$) at pH 6.5.

55 [0041] *Archaeoglobus venificus* SNP6 was isolated in the Middle Atlantic Ridge and grows optimally at 75°C ($T_{max} = 92^\circ\text{C}$) at pH 6.9.

[0042] *Aquifex pyrophilus* K01 5a was isolated at Kolbeinsey Ridge, North of Iceland. This marine organism is a gram-negative, rod-shaped, strictly chemolithoautrophic, knall gas bacterium. It grows optimally at 85°C ($T_{max} = 95^\circ\text{C}$) at pH 6.8.

[0043] M11TL is a new species of Desulfurococcus which was isolated from Diamond Pool (formerly Jim's Black Pool) in Yellowstone. The organism grows heterotrophically by fermentation of different organic materials (sulfur is not necessary) in grape-like aggregates optimally at 85 - 88°C in a low salt medium at pH 7.0.

[0044] *Thermococcus* CL-2 was isolated in the North Cleft Segment of the Juan de Fuca Ridge from a severed alvinellid worm residing on a "black smoker" sulfide structure. This marine archaea forms pleomorphic cocci, and grows optimally at 88°C.

[0045] *Aquifex* VF5 was isolated at a beach in Vulcano, Italy. This marine organism is a gram-negative, rod-shaped, strictly chemolithoautotrophic, knall gas bacterium. It grows optimally at 85°C ($T_{max} = 95^{\circ}\text{C}$) at pH 6.8.

[0046] *Teredinibacter* (pure) is an endosymbiont of the shipworm *Bankia gouldi*. The organism has straight to slightly bent 5-10 μm rods, and forms spiral cells as stationary phase is met. The organism was described in Science (1983) 22:1401-1403. It grows optimally at 30°C at pH 8.0.

[0047] *Archaeoglobus fulgidus* VC16 was isolated in Vulcano, Italy. The organism grows optimally at 85°C ($T_{max} = 92^{\circ}\text{C}$) at pH 7.0.

[0048] *Sulfolobus solfataricus* P1 grows optimally at 85°C ($T_{max} = 87^{\circ}\text{C}$) at pH 2.0.

[0049] Accordingly, the polynucleotides and enzymes encoded thereby are identified by the organism from which they were isolated, and are sometimes hereinafter referred to as F1/12LC (Figure 1 and SEQ ID NOS:23 and 33), TAG11/17LC (Figure 2 and SEQ ID NOS:24 and 34), SNP6/24LC (Figure 3 and SEQ ID NOS:25 and 35), AqP/28LC (Figure 4 and SEQ ID NOS:26 and 36), M11TL/29L (Figure 5 and SEQ ID NOS:27 and 37), CL-2/30LC (Figure 6 and SEQ ID NOS:28 and 38), VF5/34LC (Figure 7 and SEQ ID NOS:29 and 39), Trb/42L (Figure 8 and SEQ ID NOS:30 and 40), VC16/16MC (Figure 9 and SEQ ID NOS:31 and 41) and P1/8LC (Figure 10 and SEQ ID NOS: 32 and 42).

[0050] The polynucleotides and polypeptides of the present invention show identity at the nucleotide and protein level to known genes and proteins encoded thereby as shown in Table 1.

Table 1

Enzyme	Gene w/closest Homology (Organism)	Protein Similarity (%)	Protein Identity (%)	DNA Identity (%)
F1/12LC	No significant homology	-	-	-
TAG11/17LC	No significant homology	-	-	-
SNP6/24LC	PIR S34609 - carboxylesterase <i>Pseudomonas sp.</i> (strain KWI- 56) open reading frame of unknown function in <i>E.coli</i> .	46	27	42
AqP/29LC		53	31	38
M11TL/29LC	No significant homology	-	-	-
CL02/30LC	No significant homology	-	-	-
VF5/34LC	Identified by homology to 28LC; also homologous to ORF of unknown function 5' of tgs in <i>E. coli</i>	84	71	71
Trb/42L	No significant homology	-	-	-
P1-8LC				
VC16-16MC				

[0051] All the clones identified in Table 1 encode polypeptides which have esterase activity.

[0052] This invention, in addition to the isolated nucleic acid molecules encoding the enzymes of the present invention, also provides substantially similar sequences. Isolated nucleic acid sequences are substantially similar if: (i) they are capable of hybridizing under conditions hereinafter described, to the polynucleotides of SEQ ID NOS:23-32 and 43-52; (ii) or they encode DNA sequences which are degenerate to the polynucleotides of SEQ ID NOS:23-32 and 43 to 52. Degenerate DNA sequences encode the amino acid sequences of SEQ ID NOS:33-42 and 53 to 62, but have variations in the nucleotide coding sequences. As used herein, substantially similar refers to the sequences having similar identity to the sequences of the instant invention. The nucleotide sequences that are substantially the same can be identified by hybridization or by sequence comparison. Enzyme sequences that are substantially the same can

be identified by one or more of the following: proteolytic digestion, gel electrophoresis and/or microsequencing.

[0053] One means for isolating the nucleic acid molecules encoding the enzymes of the present invention is to probe a gene library with a natural or artificially designed probe using art recognized procedures (see, for example: Current Protocols in Molecular Biology, Ausubel F.M. et al. (EDS.) Green Publishing Company Assoc. and John Wiley Interscience, New York, 1989, 1992). It is appreciated by one skilled in the art that the polynucleotides of SEQ ID NOS: 23-32 and 43 to 52 or fragments thereof (comprising at least 12 contiguous nucleotides), are particularly useful probes. Other particularly useful probes for this purpose are hybridizable fragments of the sequences of SEQ ID NOS:1-22 (*i.e.*, comprising at least 12 contiguous nucleotides).

[0054] With respect to nucleic acid sequences which hybridize to specific nucleic acid sequences disclosed herein, hybridization may be carried out under conditions of reduced stringency, medium stringency or even stringent conditions. As an example of oligonucleotide hybridization, a polymer membrane containing immobilized denatured nucleic acids is first prehybridized for 30 minutes at 45°C in a solution consisting of 0.9 M NaCl, 50 mM NaH₂PO₄, pH 7.0, 5.0 mM Na₂EDTA, 0.5% SDS, 10X Denhardt's, and 0.5 mg/mL polyriboadenylic acid. Approximately 2 X 10⁷ cpm (specific activity 4-9 X 10⁸ cpm/ug) of ³²P end-labeled oligonucleotide probe are then added to the solution. After 12-16 hours of incubation, the membrane is washed for 30 minutes at room temperature in 1X SET (150 mM NaCl, 20 mM Tris hydrochloride, pH 7.8, 1 mM Na₂EDTA) containing 0.5% SDS, followed by a 30 minute wash in fresh 1X SET at Tm -10°C for the oligo-nucleotide probe. The membrane is then exposed to autoradiographic film for detection of hybridization signals.

[0055] Stringent conditions means hybridization will occur only if there is at least 90% identity, preferably at least 95% identity and most preferably at least 97% identity between the sequences. See J. Sambrook et al., *Molecular Cloning, A Laboratory Manual, 2d Ed.*, Cold Spring Harbor Laboratory (1989) which is hereby incorporated by reference in its entirety.

[0056] As used herein, a first DNA (RNA) sequence is at least 70% and preferably at least 80% identical to another DNA (RNA) sequence if there is at least 70% and preferably at least a 80% or 90% identity, respectively, between the bases of the first sequence and the bases of the another sequence, when properly aligned with each other, for example when aligned by BLASTN.

[0057] The present invention relates to polynucleotides which differ from the reference polynucleotide such that the changes are silent changes, for example the change do not alter the amino acid sequence encoded by the polynucleotide. The present invention also relates to nucleotide changes which result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference polynucleotide. In a preferred aspect of the invention these polypeptides retain the same biological action as the polypeptide encoded by the reference polynucleotide.

[0058] The polynucleotides of this invention were recovered from genomic gene libraries from the organisms listed in Table 1. Gene libraries were generated in the Lambda ZAP II cloning vector (Stratagene Cloning Systems). Mass excisions were performed on these libraries to generate libraries in the pBluescript phagemid. Libraries were generated and excisions were performed according to the protocols/methods hereinafter described.

[0059] The polynucleotides of the present invention may be in the form of RNA or DNA which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA may be double-stranded or single-stranded, and if single stranded may be the coding strand or non-coding (anti-sense) strand. The coding sequences which encodes the mature enzymes may be identical to the coding sequences shown in Figures 1-20 (SEQ ID NOS:23-32) and 43 to 52 or may be a different coding sequence which coding sequence, as a result of the redundancy or degeneracy of the genetic code, and 43-52, encodes the same mature enzymes as the DNA of Figures 1-20 (SEQ ID NOS:23-32 and 43-52).

[0060] The polynucleotide which encodes for the mature enzyme of Figures 1-20 (SEQ ID NOS:33-42) and 53-62 may include, but is not limited to: only the coding sequence for the mature enzyme; the coding sequence for the mature enzyme and additional coding sequence such as a leader sequence or a proprotein sequence; the coding sequence for the mature enzyme (and optionally additional coding sequence) and non-coding sequence, such as introns or non-coding sequence 5' and/or 3' of the coding sequence for the mature enzyme.

[0061] Thus, the term "polynucleotide encoding an enzyme (protein)" encompasses a polynucleotide which includes only coding sequence for the enzyme as well as a polynucleotide which includes additional coding and/or non-coding sequence.

[0062] The present invention further relates to variants of the hereinabove described polynucleotides which encode for fragments, analogs and derivatives of the enzymes having the deduced amino acid sequences of Figures 1-20 (SEQ ID NOS:33-42 and 53-62). The variant of the polynucleotide may be a naturally occurring allelic variant of the polynucleotide or a non-naturally occurring variant of the polynucleotide.

[0063] Thus, the present invention includes polynucleotides encoding the same mature enzymes as shown in Figures 1-20 (SEQ ID NOS:23-32 and 43-52) as well as variants of such polynucleotides which variants encode for a fragment, derivative or analog of the enzymes of Figures 1-20 (SEQ ID NOS:23-32 and 43 to 52). Such nucleotide variants include deletion variants, substitution variants and addition or insertion variants.

[0064] As hereinabove indicated, the polynucleotides may have a coding sequence which is a naturally occurring allelic variant of the coding sequences shown in Figures 1-20 (SEQ ID NOS:23-32 and 43 to 52). As known in the art, an allelic variant is an alternate form of a polynucleotide sequence which may have a substitution, deletion or addition of one or more nucleotides, which does not substantially alter the function of the encoded enzyme.

5 [0065] Fragments of the full length gene of the present invention may be used as hybridization probes for a cDNA or a genomic library to isolate the full length DNA and to isolate other DNAs which have a high sequence similarity to the gene or similar biological activity. Probes of this type preferably have at least 10, preferably at least 15, and even more preferably at least 30 bases and may contain, for example, at least 50 or more bases. The probe may also be used to identify a DNA clone corresponding to a full length transcript and a genomic clone or clones that contain the complete gene including regulatory and promoter regions, exons and introns. An example of a screen comprises isolating the coding region of the gene by using the known DNA sequence to synthesize an oligonucleotide probe. Labeled oligonucleotides having a sequence complementary to that of the gene of the present invention are used to screen a library of genomic DNA to determine which members of the library the probe hybridizes to.

10 [0066] It is also appreciated that such probes can be and are preferably labeled with an analytically detectable reagent to facilitate identification of the probe. Useful reagents include but are not limited to radioactivity, fluorescent dyes or enzymes capable of catalyzing the formation of a detectable product. The probes are thus useful to isolate complementary copies of DNA from other sources or to screen such sources for related sequences.

15 [0067] The present invention further relates to polynucleotides or fragments thereof encoding an enzyme having esterase activity and which hybridize to the hereinabove-described sequences if there is at least 70%, preferably at least 90%, and more preferably at least 95% identity between the sequences. The present invention particularly relates to polynucleotides which hybridize under stringent conditions to the hereinabove-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 95 % and preferably at least 97 % identity between the sequences. The polynucleotides which hybridize to the hereinabove described polynucleotides in a preferred embodiment encode enzymes which either retain substantially the same biological function or 20 activity as the mature enzyme encoded by the DNA of Figures 1-20 (SEQ ID NOS:23-32 and 43 to 52.)

25 [0068] Alternatively, the polynucleotide may have at least 15 bases, preferably at least 30 bases, and more preferably at least 50 bases of a polynucleotide of the present invention and which has an identity thereto, as hereinabove described, and which may or may not retain activity. For example, such polynucleotides may be employed as probes for the polynucleotides of SEQ ID NOS:23-32, and 43 to 52 for example, for recovery of the polynucleotide or as a diagnostic probe or as a PCR primer.

30 [0069] Thus, the present invention is directed to polynucleotides or fragments thereof encoding an enzyme having esterase activity and having at least a 70% identity, preferably at least 90% identity and more preferably at least a 95% identity to a polynucleotide which encodes the enzymes of SEQ ID NOS:33-42 and 53 to 62 as well as fragments thereof, which fragments have at least 15 bases, preferably at least 30 bases and most preferably at least 50 bases. 35 which fragments are at least 90% identical, preferably at least 95 % identical and most preferably at least 97% identical under stringent conditions to any portion of a polynucleotide of the present invention.

[0070] The present invention further relates to enzymes which have the deduced amino acid sequences of Figures 1-20 (SEQ ID NOS:23-32 and 43 to 52) as well as fragments, analogs and derivatives of such enzyme.

40 [0071] The terms "fragment," "derivative" and "analog" when referring to the enzymes of Figures 1-20 (SEQ ID NOS: 33-42 and 53 to 62) mean enzymes which retain essentially the same biological function or activity as such enzymes. Thus, an analog includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature enzyme.

[0072] The enzymes of the present invention may be a recombinant enzyme, a natural enzyme or a synthetic enzyme, preferably a recombinant enzyme.

45 [0073] The fragment, derivative or analog of the enzymes of Figures 1-20 (SEQ ID NOS:33-42 and 53 to 62) may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature enzyme is fused with another compound, such as a compound to increase the half-life of the enzyme (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature enzyme, such as a leader or secretory sequence or a sequence which is employed for purification of the mature enzyme or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

50 [0074] The enzymes and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

[0075] The term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or enzyme present in a living animal is not isolated, but the same polynucleotide or enzyme, separated from some or all of the coexisting materials

in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or enzymes could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment.

[0076] The enzymes of the present invention include the enzymes of SEQ ID NOS:33-42 and 53 to 62 (in particular the mature enzyme) as well as enzymes which have at least 70 % similarity (preferably at least 70% identity) to the enzymes of SEQ ID NOS:33-42 and 53 to 62 and more preferably at least 90% similarity (more preferably at least 90% identity) to the enzymes of SEQ ID NOS:33-42 and 53 to 62 and still more preferably at least 95% similarity (still more preferably at least 95% identity) to the enzymes of SEQ ID NOS:33-42 and 53 to 62 and also include portions of such enzymes with such portion of the enzyme generally containing at least 30 amino acids and more preferably at least 50 amino acids.

[0077] As known in the art "similarity" between two enzymes is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one enzyme to the sequence of a second enzyme.

[0078] A variant, i.e. a "fragment", "analog" or "derivative" polypeptide, and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions, fusions and truncations, which may be present in any combination.

[0079] Among preferred variants are those that vary from a reference by conservative amino acid substitutions. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr.

[0080] Most highly preferred are variants which retain the same biological function and activity as the reference polypeptide from which it varies.

[0081] Fragments or portions of the enzymes of the present invention may be employed for producing the corresponding full-length enzyme by peptide synthesis; therefore, the fragments may be employed as intermediates for producing the full-length enzymes. Fragments or portions of the polynucleotides of the present invention may be used to synthesize full-length polynucleotides of the present invention.

[0082] The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of enzymes of the invention by recombinant techniques.

[0083] Host cells are genetically engineered (transduced or transformed or transfected) with the vectors of this invention which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the genes of the present invention. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

[0084] The polynucleotides of the present invention may be employed for producing enzymes by recombinant techniques. Thus, for example, the polynucleotide may be included in any one of a variety of expression vectors for expressing an enzyme. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host.

[0085] The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art.

[0086] The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the *E. coli*. lac or trp, the phage lambda P_L promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

[0087] In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in *E. coli*.

[0088] The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein.

[0089] As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as *E. coli*.

Streptomyces, *Bacillus subtilis*; fungal cells, such as yeast; insect cells such as *Drosophila S2* and *Spodoptera Sf9*; animal cells such as CHO, COS or Bowes melanoma; adenoviruses; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

[0090] More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example; Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pBluescript II KS, pTRC99a, pKK223-3, pDR540, pRIT2T (Pharmacia); Eukaryotic: pXT1, pSG5 (Stratagene) pSVK3, pBPV, pMSG, pSVL, SV40 (Pharmacia). However, any other plasmid or vector may be used as long as they are replicable and viable in the host.

[0091] Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK223-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda P_R, P_L and trp. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

[0092] In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (Davis, L., Dibner, M., Battey, J., *Basic Methods in Molecular Biology*, (1986)).

[0093] The constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the enzymes of the invention can be synthetically produced by conventional peptide synthesizers.

[0094] Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual, Second Edition*, Cold Spring Harbor, N.Y., (1989), the disclosure of which is hereby incorporated by reference.

[0095] Transcription of the DNA encoding the enzymes of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act on a promoter to increase its transcription. Examples include the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

[0096] Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), α-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated enzyme. Optionally, the heterologous sequence can encode a fusion enzyme including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

[0097] Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

[0098] As a representative but nonlimiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

[0099] Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period.

[0100] Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

[0101] Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods are well known to those skilled in the art.

[0102] Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, *Cell*, 23: 175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

[0103] The enzyme can be recovered and purified from recombinant cell cultures by methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

[0104] The enzymes of the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the enzymes of the present invention may be glycosylated or may be non-glycosylated. Enzymes of the invention may or may not also include an initial methionine amino acid residue.

[0105] Esterases are a group of key enzymes in the metabolism of fats and are found in all organisms from microbes to mammals. In the hydrolysis reaction, an ester group is hydrolysed to an organic acid and an alcohol.

[0106] Esterases enantiomerically differentiate dicarboxylic diesters and diacetates of diols. Using the approach disclosed in a commonly assigned, copending provisional application Serial No. 60/008,316, filed on December 7, 1995 and entitled "Combinatorial Enzyme Development," the disclosure of which is incorporated herein by reference in its entirety, one could convert the enantiospecificity of the esterase. Further, the thermostable esterases are believed to have superior stability at higher temperatures and in organic solvents. Thus, they are better suited for use in rigorous production processes which require robust catalysts.

[0107] There are a number of industrial and scientific applications for esterases, such as those of the present invention, including:

1) Esterases are useful in the dairy industry as ripening starters for cheeses, such as the Swiss-type cheeses;

2) Esterases are useful in the pulp and paper industry for lignin removal from cellulose pulps, for lignin solubilization by cleaving the ester linkages between aromatic acids and lignin and between lignin and hemicelluloses, and for disruption of cell wall structure when used in combination with xylanase and other xylan-degrading enzymes in biopulping and biobleaching of pulps;

3) Esterases are useful in the synthesis of carbohydrate derivatives, such as sugar derivatives;

4) Esterases are useful, when combined with xylanases and cellulases, in the conversion of lignocellulosic wastes to fermentable sugars for producing a variety of chemicals and fuels;

5) Esterases are useful as research reagents in studies on plant cell wall structure, particularly the nature of covalent bonds between lignin and carbohydrate polymers in the cell wall matrix;

6) Esterases are also useful as research reagents in studies on mechanisms related to disease resistance in plants and the process of organic matter decomposition; and

7) Esterases are useful in selection of plants bred for production of highly digestible animal feeds, particularly for ruminant animals.

[0108] Antibodies generated against the enzymes corresponding to a sequence of the present invention can be obtained by direct injection of the enzymes into an animal or by administering the enzymes to an animal, preferably a nonhuman. The antibody so obtained will then bind the enzymes itself. In this manner, even a sequence encoding only a fragment of the enzymes can be used to generate antibodies binding the whole native enzymes. Such antibodies

can then be used to isolate the enzyme from cells expressing that enzyme.

[0109] For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler and Milstein, *Nature*, 256:495-497, 1975), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, *Immunology Today* 4:72, 1983), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole *et al.*, in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96, 1985).

[0110] Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce single chain antibodies to immunogenic enzyme products of this invention. Also, transgenic mice may be used to express humanized antibodies to immunogenic enzyme products of this invention.

[0111] Antibodies generated against an enzyme of the present invention may be used in screening for similar enzymes from other organisms and samples. Such screening techniques are known in the art, for example, one such screening assay is described in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual* (2d Ed.), Cold Spring Harbor Laboratory, Section 12.21-12.28 (1989) which is hereby incorporated by reference in its entirety.

[0112] The present invention will be further described with reference to the following examples; however, it is to be understood that the present invention is not limited to such examples. All parts or amounts, unless otherwise specified, are by weight.

[0113] In order to facilitate understanding of the following examples certain frequently occurring methods and/or terms will be described.

[0114] "Plasmids" are designated by a lower case "p" preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

[0115] "Digestion" of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors and other requirements were used as would be known to the ordinarily skilled artisan. For analytical purposes, typically 1 µg of plasmid or DNA fragment is used with about 2 units of enzyme in about 20 µl of buffer solution. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 µg of DNA are digested with 20 to 250 units of enzyme in a larger volume. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37° C are ordinarily used, but may vary in accordance with the supplier's instructions. After digestion the reaction is electrophoresed directly on a polyacrylamide gel to isolate the desired fragment.

[0116] Size separation of the cleaved fragments is performed using 8 percent polyacrylamide gel described by Goeddel *et al.*, *Nucleic Acids Res.*, 8:4057 (1980).

[0117] "Oligonucleotides" refers to either a single stranded polydeoxynucleotide or two complementary polydeoxy-nucleotide strands which may be chemically synthesized. Such synthetic oligonucleotides have no 5' phosphate and thus will not ligate to another oligonucleotide without adding a phosphate with an ATP in the presence of a kinase. A synthetic oligonucleotide will ligate to a fragment that has not been dephosphorylated.

[0118] "Ligation" refers to the process of forming phosphodiester bonds between two double stranded nucleic acid fragments (Maniatis, T., *et al.*, *Id.*, p. 146). Unless otherwise provided, ligation may be accomplished using known buffers and conditions with 10 units of T4 DNA ligase ("ligase") per 0.5 µg of approximately equimolar amounts of the DNA fragments to be ligated.

[0119] Unless otherwise stated, transformation was performed as described in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual* (2d Ed.), Cold Spring Harbor Press (1989).

Example 1

Bacterial Expression and Purification of Esterases

[0120] DNA encoding the enzymes of the present invention, SEQ ID NOS:33 through 42, were initially amplified from a pBluescript vector containing the DNA by the PCR technique using the primers noted herein. The amplified sequences were then inserted into the respective PQE vector listed beneath the primer sequences, and the enzyme was expressed according to the protocols set forth herein. The 5' and 3' primer sequences for the respective genes are as follows:

Staphylothermus marinus F1-12LC

5' CCGAGAATTTC ATTAAAGAGG AGAAATTAAC TATGTCTTTAACAAAGCACT CT
3' CGGAAGATCT CTATCGTTAAC GTGTATGATT T

5

vector: pQET

Pyrodictium TAG 11-17LC

10

5' CCGAGAATTTC ATTAAAGAGG AGAAATTAAC TATGAAACTC CTTGAGGCCA CA EcoRI
3' CGGAAGATCT CGCCGGTACA CCATCAGCCA C BgIII

15

vector: pQET

Archaeoglobus venificus SNP6-24LC

20

5' CCGAGAATTTC ATTAAAGAGG AGAAATTAAC TATGCCATAT GTTAGGAATG GT
3' CGGAGGTACC TTAGAACTGT GCTGAAGAAA TAAATTGTC CATTGCTCT
3' CGGAGGTACC TTAGAACTGT GCTGAAGAAA TAAATTGTC CATTGCTCTA TTA

25

vector: pQET

Aquifex pyrophilus - 28LC

30

5' CCGAGAATTTC ATTAAAGAGG AGAAATTAAC TATGAGATTG AGGAAATTG AAG
3' CGGAGGTACC CTATTCAAGAA AGTACCTCTA A

35

vector: pQET

M11TL - 29LC

40

5' CCGAGAATTTC ATTAAAGAGG AGAAATTAAC TATGTAAAT ATCAATGTCT TT
3' CGGAAGATCT TTAAGGATTTC CCGCTGGGTA G

45

vector: pQET

Thermococcus CL-2 - 30LC

50

5' CCGAGAATTTC ATTAAAGAGG AGAAATTAAC TATGGAGGTT TACAAGGCCA AA
3' CGGAGGTACC TTATTGAGCC GAAGAGTACG A

55

vector: pQET

Aquifex VF5 - 34LC

5' CCGAGAATTTC ATAAAGAGG AGAAATTAAC TATGATTGGC AATTTGAAT TGA EcoRI
 3' CGGAGGTACC TTAAAGTGCT CTCATATCCC C KpnI

5

vector: pQET

Teredinibacter 42L

10

5' CCGAGAATTTC ATAAAGAGG AGAAATTAAC TATGCCAGCT AATGACTCAC CC
 3' CGGAAGATCT TCAACAGGCT CCAAATAATT TC (without His-tag)
 3' CGGAAGATCT ACAGGCTCCA AATAATTTC (with His-tag)

15

vector: pQE12

Archaeoglobus fulgidus VC16-16MC

20

5' CCGAGAATTTC ATAAAGAGG AGAAATTAAC TATGCTTGAT ATGCCAATCG AC EcoRI
 3' CGGAGGTACC CTAGTCGAAG ACAAGAAGAG C KpnI

25

vector: pQET

Sulfolabrus solfataricus P1-8LC

30

5' CCGAGAATTTC ATAAAGAGG AGAAATTAAC TATGCCAG GATCCTAGAA TT EcoRI
 3' CGGAGGTACC TTAAATTTA TCATAAAAATA C KpnI

vector: pQET

35 [0121] The restriction enzyme sites indicated correspond to the restriction enzyme sites on the bacterial expression vector indicated for the respective gene (Qiagen, Inc. Chatsworth, CA). The pQE vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter operator (P/O), a ribosome binding site (RBS), a 6-His tag and restriction enzyme sites.

40 [0122] The pQE vector was digested with the restriction enzymes indicated. The amplified sequences were ligated into the respective pQE vector and inserted in frame with the sequence encoding for the RBS. The ligation mixture was then used to transform the *E. coli* strain M15/pREP4 (Qiagen, Inc.) by electroporation. M15/pREP4 contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants were identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies were selected. Plasmid DNA was isolated and confirmed by restriction analysis. Clones containing the desired constructs were grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 µg/ml) and Kan (25 µg/ml). The O/N culture was used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells were grown to an optical density 600 (O.D.₆₀₀) of between 0.4 and 0.6. IPTG ("Isopropyl-B-D-thiogalacto pyranoside") was then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression. Cells were grown an extra 3 to 4 hours. Cells were then harvested by centrifugation.

50 [0123] The primer sequences set out above may also be employed to isolate the target gene from the deposited material by hybridization techniques described above.

Example 2

Isolation of a Selected Clone from the Deposited Genomic Clones

[0124] The two oligonucleotide primers corresponding to the gene of interest are used to amplify the gene from the deposited material. A polymerase chain reaction is carried out in 25 µl of reaction mixture with 0.1 µg of the DNA of

the gene of interest. The reaction mixture is 1.5-5 mM MgCl₂, 0.01 % (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 1.25 Unit of Taq polymerase. Thirty cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with the Perkin-Elmer Cetus 9600 thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the gene of interest by subcloning and sequencing the DNA product.

Example 3

10 Production of the Expression Gene Bank

[0125] Colonies containing pBluescript plasmids with random inserts from the organisms M11TL, *Thermococcus* GU5L5, and *Teredinibacter* were obtained according to the method of Hay and Short, *Strategies*, 5:16, 1992.

15 Example 4

Screening for Lipase/Esterase Activity

[0126] The resulting colonies were picked with sterile toothpicks and used to singly inoculate each of the wells of 96-well microtiter plates. The wells contained 250 µL of LB media with 100 µg/mL ampicillin, 80 µg/mL methicillin, and 10% v/v glycerol (LB Amp/Meth, glycerol). The cells were grown overnight at 37°C without shaking. This constituted generation of the "Source GeneBank." Each well of the Source GeneBank thus contained a stock culture of *E. coli* cells, each of which contained a pBluescript with a unique DNA insert.

[0127] The plates of the Source GeneBank were used to multiply inoculate a single plate (the "Condensed Plate") containing in each well 200 µL of LB Amp/Meth, glycerol. This step was performed using the High Density Replicating Tool (HDRT) of the Beckman Biomek with a 1 % bleach, water, isopropanol, air-dry sterilization cycle in between each inoculation. Each well of the Condensed Plate thus contained 10 to 12 different pBluescript clones from each of the source library plates. The Condensed Plate was grown for 16 hours at 37°C and then used to inoculate two white 96-well Polyfiltronics microtiter daughter plates containing in each well 250 µL of LB Amp/Meth (no glycerol). The original condensed plate was put in storage -80°C. The two condensed daughter plates were incubated at 37°C for 18 hours.

[0128] The short chain esterase '600 µM substrate stock solution' was prepared as follows: 25 mg of each of the following compounds was dissolved in the appropriate volume of DMSO to yield a 25.2 mM solution. The compounds used were 4-methylumbelliferyl propionate, 4-methylumbelliferyl butyrate, and 4-methylumbelliferyl heptanoate. Two hundred fifty microliters of each DMSO solution was added to ca 9 mL of 50 mM, pH 7.5 Hepes buffer which contained 0.6% of Triton X-100 and 0.6 mg per mL of dodecyl maltoside (Anatrace). The volume was taken to 10.5 mL with the above Hepes buffer to yield a slightly cloudy suspension.

[0129] The long chain '600 µM substrate stock solution' was prepared as follows: 25 mg of each of the following compounds was dissolved in DMSO to 25.2 mM as above. The compounds used were 4-methylumbelliferyl elaidate, 4-methylumbelliferyl palmitate, 4-methylumbelliferyl oleate, and 4-methylumbelliferyl stearate. All required brief warming in a 70°C bath to achieve dissolution. Two hundred fifty microliters of each DMSO solution was added to the Hepes buffer and diluted to 10.5 mL as above. All seven umbelliferones were obtained from Sigma Chemical Co.

[0130] Fifty µL of the long chain esterase or short chain esterase '600 µM substrate stock solution' was added to each of the wells of a white condensed plate using the Biomek to yield a final concentration of substrate of about 100 µM. The fluorescence values were recorded (excitation = 326 nm, emission = 450 nm) on a plate-reading fluorometer immediately after addition of the substrate. The plate was incubated at 70°C for 60 minutes in the case of the long chain substrates, and 30 minutes at RT in the case of the short chain substrates. The fluorescence values were recorded again. The initial and final fluorescence values were compared to determine if an active clone was present.

50 Example 5

Isolation and Purification of the Active Clone

[0131] To isolate the individual clone which carried the activity, the Source GeneBank plates were thawed and the individual wells used to singly inoculate a new plate containing LB Amp/Meth. As above, the plate was incubated at 37°C to grow the cells, 50 µL of 600 µM substrate stock solution was added using the Biomek and the fluorescence was determined. Once the active well from the source plate was identified, cells from this active well were streaked on agar with LB/Amp/Meth and grown overnight at 37°C to obtain single colonies. Eight single colonies were picked with

a sterile toothpick and used to singly inoculate the wells of a 96-well microtiter plate. The wells contained 250 µL of LB Amp/Meth. The cells were grown overnight at 37°C without shaking. A 200 µL aliquot was removed from each well and assayed with the appropriate long or short chain substrates as above. The most active clone was identified and the remaining 50 µL of culture was used to streak an agar plate with LB/Amp/Meth. Eight single colonies were picked,
5 grown and assayed as above. The most active clone was used to inoculate 3 mL cultures of LB/Amp/Meth, which were grown overnight. The plasmid DNA was isolated from the cultures and utilized for sequencing.

SEQUENCE LISTING

10 [0132]

(1) GENERAL INFORMATION:

(i) APPLICANTS:

15 ROBERTSON, Daniel E.

MURPHY, Dennis

20 REID, John

MAFFIA, Anthony

LINK, Steven

25 SWANSON, Ronald V.

WARREN, Patrick V.

30 KOSMOTKA, Anna

CALLEN, Walter

(ii) TITLE OF INVENTION:

35 ESTERASES

(iii) NUMBER OF SEQUENCES: 62

40 (iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI, STEWART & OLSTEIN
(B) STREET: 6 BECKER FARM ROAD
(C) CITY: ROSELAND
45 (D) STATE: NEW JERSEY
(E) COUNTRY: USA
(F) ZIP: 07068

50 (v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: 3.5 INCH DISKETTE
(B) COMPUTER: IBM PS/2
(C) OPERATING SYSTEM: MS-DOS
(D) SOFTWARE: WORD PERFECT 5.1

55 (vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: Unassigned

EP 0 880 590 B1

(B) FILING DATE: Concurrently
(C) CLASSIFICATION: Unassigned

5 (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

10 (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: HERRON, CHARLES J.
(B) REGISTRATION NUMBER: 28,019
(C) REFERENCE/DOCKET NUMBER: 331400-39

15 (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 201-994-1700
(B) TELEFAX: 201-994-1744

20 (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS

25 (A) LENGTH: 52 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

30 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

35 **CCGAGAAATTCA TTAAAGAGGG AGAAATTAAC TATGTCTTTA ACAAGCACT CT**

52

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS

40 (A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

45 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

50 **CGGAAGATCT CTATCGTTTA GTGTATGATT T**

31

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 52 NUCLBOTIDES
(B) TYPE: NUCLEIC ACID

EP 0 880 590 B1

- (C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCGAGAATTTC ATTAAAGAGG AGAAATTAAC TATGAAACTC CTTGAGCCCCA CA

52

10

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS

15

- (A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

20

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

25

CGGAAGATCT CGCCGGTACA CCATCAGCCA C

31

30

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS

35

- (A) LENGTH: 52 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

35

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

40

CCGAGAATTTC ATTAAAGAGG AGAAATTAAC TATGCCATAT GTTAGGAATG GT

52

45

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS

50

- (A) LENGTH: 53 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

55

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CGGAGGTACC TTAGAACTGT GCTGAAGAAA TAAATTCGTC CATTGCTCTA TTA

53

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS

- 5 (A) LENGTH: 49 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

10 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

15 CGGAGGTACC TTAGAACCTGT GCTGAAGAAA TAAATTCGTC CATTGCTCT.

49

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS

- 20 (A) LENGTH: 53 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

25 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

30 CCGAGAAATTTC ATTAAAGAGG AGAAATTAAC TATGAGATTG AGGAAATTG AAG

53

(2) INFORMATION FOR SEQ ID NO:9:

35 (i) SEQUENCE CHARACTERISTICS

- 40 (A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

45 CGGAGGTACC CTATTCAGAA AGTACCTCTA A

31

(2) INFORMATION FOR SEQ ID NO:10:

50 (i) SEQUENCE CHARACTERISTICS

- 55 (A) LENGTH: 52 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCGAGAATTTC ATTAAAGAGG AGAAATTAAC TATGTTTAAT ATCAATGTCT TT

52

5

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS

10

- (A) LENGTH: 31 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

15

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

20

CGGAAGATCT TTAAGGATT TCCCTGGGTA G

31

25

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS

25

- (A) LENGTH: 52 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

30

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

35

CCGAGAATTTC ATTAAAGAGG AGAAATTAAC TATGGAGGTT TACAAGGCCA AA

52

35

(2) INFORMATION FOR SEQ ID NO:13:

40

(i) SEQUENCE CHARACTERISTICS

40

- (A) LENGTH: 31 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

45

(ii) MOLECULE TYPE: cDNA

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGAGGTACC TTATTGAGCC GAAGAGTACG A

31

55

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 53 NUCLEOTIDES

EP 0 880 590 B1

- (B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

5 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

10 **CCGAGAATTTC ATAAAGAGG AGAAATTAAC TATGATTGGC AATTTGAAT TGA**

53

(2) INFORMATION FOR SEQ ID NO:15:

15 (i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

20 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

25 **CGGAGGTACC TTAAAGTGCT CTCATATCCC C**

31

(2) INFORMATION FOR SEQ ID NO:16:

30 (i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

35 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

40 **CGGAGAATTTC ATAAAGAGG AGAAATTAAC TATGCCAGCT AATGACTCAC CC**

52

(2) INFORMATION FOR SEQ ID NO:17:

45 (i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 32 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

50 (ii) MOLECULE TYPE: cDNA

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGGAAGATCT TCAACAGGCT CCAAATAATT TC

32

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS

- 5 (A) LENGTH: 29 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

10 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

15 CGGAAGATCT ACAGGGCTCCA AATAATTTC

29

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS

- 20 (A) LENGTH: 52 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

25 (ii) MOLECULE TYPE : cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

30 CCGAGAAATTTC ATTAAGAGGG AGAAATTAAAC TATGCTTGAT ATGCCAATCG AC

52

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS

- 35 (A) LENGTH: 31 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

40 (ii) MOLECULE TYPE: cDNA

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGGAGGTACC CTAGTCGAAC AGAAGAAGAG C

31

50 (2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS

- 55 (A) LENGTH: 52 NUCLEOTIDES
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

5

CCGAGAATTC ATAAAGAGG AGAAATTAAAC TATGCCCTTA GATCCTAGAA TT

52

(2) INFORMATION FOR SEQ ID NO:22:

10

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 31 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

15

(ii) MOLECULE TYPE: cDNA

20

CGGAGGTACC TAAATTTTA TCATAAAATA C

31

(2) INFORMATION FOR SEQ ID NO:23:

25

(i) SEQUENCE CHARACTERISTICS

30

- (A) LENGTH: 555 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

40

45

50

55

EP 0 880 590 B1

ATG TCT TTA AAC AAG CAC TCT TGG ATG GAT ATG ATA ATA TTT ATT CTC Met Ser Leu Asn Lys His Ser Trp Met Asp Met Ile Ile Phe Ile Leu 1 5 10 15	48
5 AGC TTT TCT TTC CCA TTA ACA ATG ATC GCA TTA GCT ATC TCT ATG TCG Ser Phe Ser Phe Pro Leu Thr Met Ile Ala Leu Ala Ile Ser Met Ser 20 25 30	96
10 TCA TGG TTT AAT ATA TGG AAT ATA GCA TTA AGC GAT CTA CGA CAT GCT Ser Trp Phe Asn Ile Trp Asn Asn Ala Leu Ser Asp Leu Gly His Ala 35 40 45	144
15 GTT AAA AGC AGT GTT GCT CCA ATA TTC AAT CTA GGT CTT GCA ATT GGT Val Lys Ser Ser Val Ala Pro Ile Phe Asn Leu Gly Leu Ala Ile Gly 50 55 60	192
20 25 GGG ATA CTA ATT GTT ATA GTT GGT TTA AGA AAT CTT TAT TCG TGG AGT Gly Ile Leu Ile Val Ile Val Gly Leu Arg Asn Leu Tyr Ser Trp Ser 65 70 75 80	240
20 AGA GTT AAA GGA TCT TTA ATC ATA TCC ATG GGT GTA TTT CTT AAC TTA Arg Val Lys Gly Ser Leu Ile Ile Ser Met Gly Val Phe Leu Asn Leu 85 90 95	288
25 30 ATA GGG GTT TTC GAC GAA GTA TAT GGT TGG ATA CAT TTC CTA GTC TCA Ile Gly Val Phe Asp Glu Val Tyr Gly Trp Ile His Phe Leu Val Ser 100 105 110	336
30 GTA TTG TTT TTC TTA TCA ATA ATA GCA TAT TTC ATA GCT ATA TCA ATA Val Leu Phe Phe Leu Ser Ile Ile Ala Tyr Phe Ile Ala Ile Ser Ile 115 120 125	384
35 CTT GAC AAA TCA TGG ATA GCT GTT CTA CTA ATA ATA GGT CAT ATT GCA Leu Asp Lys Ser Trp Ile Ala Val Leu Leu Ile Ile Gly His Ile Ala 130 135 140	432
40 45 ATG TGG TAT CTA CAC TTT GCT TCA GAG ATT CCG AGA GGT GCG GCT ATT Met Trp Tyr Leu His Phe Ala Ser Glu Ile Pro Arg Gly Ala Ala Ile 145 150 155 160	480
35 CCC GAG TTA TTA GCG GTA TTC TCG TTT TTA CCA TTC TAT ATA AGA GAC Pro Glu Leu Leu Ala Val Phe Ser Phe Leu Pro Phe Tyr Ile Arg Asp 165 170 175	528
40 TAT TTT AAA TCA TAC ACT AAA CGA TAG Tyr Phe Lys Ser Tyr Thr Lys Arg 180	576

(2) INFORMATION FOR SEQ ID NO:24:

- 45 (i) SEQUENCE CHARACTERISTICS
- (A) LENGTH: 1041 NUCLEOTIDES
 - (B) TYPE: NUCLEIC ACID
 - 50 (C) STRANDEDNESS: SINGLE
 - (D) TOPOLOGY: LINEAR
- 40 (ii) MOLECULE TYPE: GENOMIC DNA
- 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

EP 0 880 590 B1

	ATG AAA CTC CTT GAG CCC ACA AAT ACC TCC TAC ACG CTG TTA CAG GAT Met Lys Leu Leu Glu Pro Thr Asn Thr Ser Tyr Thr Leu Leu Gln Asp 1 5 10 15	48
5	TAA GCA TTG CAT TTT GCA TTT TAC TGG TTT CTG GCC GTG TAT ACG TGG Leu Ala Leu His Phe Ala Phe Tyr Trp Phe Leu Ala Val Tyr Thr Trp 20 25 30	
10	TAA CCC GGT GTC CTA GTC CGG GGC GTC GCT GTG GAC ACA GGG GTG GCT Leu Pro Gly Val Leu Val Arg Gly Val Ala Val Asp Thr Gly Val Ala 35 40 45	144
	CGG GTG CCT GGG CTC GGC CGG CGC GGT AAG AGG CTG CTC CTG GCC GCT Arg Val Pro Gly Leu Gly Arg Arg Gly Lys Arg Leu Leu Ala Ala 50 55 60	192
15	GTC GCT GTC TTG GCG CTT GTT GTG TCC GTT GTT GTC CCG GCT TAT GTG Val Ala Val Leu Ala Leu Val Val Ser Val Val Val Pro Ala Tyr Val 65 70 75 80	240
20	GCG TAT AGT AGT CTG CAC CCG GAG AGC TGT CGG CCC GTT GCG CCG GAG Ala Tyr Ser Ser Leu His Pro Glu Ser Cys Arg Pro Val Ala Pro Glu 85 90 95	288
	GGG CTC ACC TAC AAA GAG TTC AGC GTG ACC GCG GAG GAT GGC TTG GTG Gly Leu Thr Tyr Lys Glu Phe Ser Val Thr Ala Glu Asp Gly Leu Val 100 105 110	336
25	GTT CGG GGC TGG GTG CTG GGC CCC GGC GCT GGG GGC AAC CCG GTG TTC Val Arg Gly Trp Cys Leu Gly Pro Gly Ala Gly Gly Asn Pro Val Phe 115 120 125	384
	GTT TTG ATG CAC GGG TAT ACT GGG TGC CGC TCG GCG CCC TAC ATG GCT Val Leu Met His Gly Tyr Thr Gly Cys Arg Ser Ala Pro Tyr Met Ala 130 135 140	432
30	GTG CTG GCC CGG GAG CTC GTG GAG TGG GGG TAC CCG GTG GTT GTG TTC Val Leu Ala Arg Glu Leu Val Glu Trp Gly Tyr Pro Val Val Val Phe 145 150 155 160	480
35	GAC TTC CGG GGC CAC GGG GAG AGC GGG GGC TCG ACG ACG ATT GGG CCC Asp Phe Arg Gly His Gly Glu Ser Gly Gly Ser Thr Thr Ile Gly Pro	528

40

45

50

55

EP 0 880 590 B1

	165	170	175	
5	CGG GAG GTG CTG GAT GCC CGG GCT GTG GTG GGC TAT GTC TCG GAG CGG Arg Glu Val Leu Asp Ala Arg Ala Val Val Gly Tyr Val Ser Glu Arg 180 185 190			576
	TTC CCC GGC CGC CCG ATA ATA TTG GTG GGG TTC AGT ATG GGC GGC GCT Phe Pro Gly Arg Arg Ile Ile Leu Val Gly Phe Ser Met Gly Gly Ala 195 200 205			624
10	GTA GCG ATC GTG GAG GGT GCT GGG GAC CCG CGG GTC TAC GCG GTG GCT Val Ala Ile Val Glu Gly Ala Gly Asp Pro Arg Val Tyr Ala Val Ala 210 215 220			672
15	GCT GAT AGC CCG TAC TAT AGG CTC CGG GAC GTC ATA CCC CGG TGG CTG Ala Asp Ser Pro Tyr Tyr Arg Leu Arg Asp Val Ile Pro Arg Trp Leu 225 230 235 240			720
	GAG TAC AAG ACG CCG CTG CCG GGC TGG GTG GGT GTG CTG GCC GGG TTC Glu Tyr Lys Thr Pro Leu Pro Gly Trp Val Gly Val Leu Ala Gly Phe 245 250 255			768
20	TAC GGG AGG CTG ATG GCG GGC GTT GAC CTC CGC TTC GGC CCC GCT GGG Tyr Gly Arg Leu Met Ala Gly Val Asp Leu Gly Phe Gly Pro Ala Gly 260 265 270			816
25	GTG GAG CGC GTG GAT AAG CCG TTG CTG GTG GTG TAT GGG CCC CGG GAC Val Glu Arg Val Asp Lys Pro Leu Leu Val Val Tyr Gly Pro Arg Asp 275 280 285			864
	CCG CTG GTG ACG CGG GAC GAG GCG AGG AGC CTG GCG TCC CGT AGC CCG Pro Leu Val Thr Arg Asp Glu Ala Arg Ser Leu Ala Ser Arg Ser Pro 290 295 300			912
30	TGT GGC CGT CTC GTC GAG GTT CCT GCG GCT GGC CAC GTG GAG GCC GTG Cys Gly Arg Leu Val Glu Val Pro Gly Ala Gly His Val Glu Ala Val 305 310 315 320			960
	GAT GTG CTC GGG CCG GGC CGC TAC GCA GAC ATG CTG ATA GAG CTG CCG Asp Val Leu Gly Pro Gly Arg Tyr Ala Asp Met Leu Ile Glu Leu Ala 325 330 335			1008
35	CAC GAG GAG TGC CCT CCG GGG GCC GGT GGC TGA His Glu Glu Cys Pro Pro Gly Ala Gly Gly 340 345			1029

40 (2) INFORMATION FOR SEO ID NO:25:

(i) SEQUENCE CHARACTERISTICS

- 45 (A) LENGTH: 789 NUCLEOTIDES
 (B) TYPE: NUCLEIC ACID
 (C) STRANDEDNESS: SINGLE
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

50

55

EP 0 880 590 B1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

	ATG CCA TAT GTT AGG AAT GGT GTC AAT ATC TAT TAT GAA CTG GTG Met Pro Tyr Val Arg Asn Gly Gly Val Asn Ile Tyr Tyr Glu Leu Val 1 5 10 15	48
5	GAT GGA CCT GAG CCA CCA ATT GTC TTT GTT CAC GGA TGG ACA GCA AAT Asp Gly Pro Glu Pro Pro Ile Val Phe Val His Gly Trp Thr Ala Asn 20 25 30	96
10	ATG AAT TTT TGG AAA GAG CAA AGA CGT TAT TTT GCA GGC AGG AAT ATG Met Asn Phe Trp Lys Glu Gln Arg Arg Tyr Phe Ala Gly Arg Asn Met 35 40 45	144
15	ATG TTG TTT GTC GAT AAC AGA GGT CAT GGC AGG TCC GAT AAG CCA CTT Met Leu Phe Val Asp Asn Arg Gly His Gly Arg Ser Asp Lys Pro Leu 50 55 60	192
20	GGA TAC GAT TTC TAC AGA TTT GAG AAC TTC ATT TCA GAT TTA GAT GCG Gly Tyr Asp Phe Tyr Arg Phe Glu Asn Phe Ile Ser Asp Leu Asp Ala 65 70 75 80	240
25	GTT GTT AGG GAG ACT GGA GTG GAG AAA TTT GTC CTC GTC GGA CAT TCA Val Val Arg Glu Thr Gly Val Glu Lys Phe Val Leu Val Gly His Ser 85 90 95	288
30	TTC GGA ACA ATG ATC TCT ATG AAG TAC TGT TCG GAG TAT CGG AAT CCG Phe Gly Thr Met Ile Ser Met Lys Tyr Cys Ser Glu Tyr Arg Asn Arg 100 105 110	336
35	GTT CTT GCT CTA ATC CTC ATA GGT GGT GGG AGC AGA ATA AAG CTT CTA Val Leu Ala Ile Leu Ile Gly Gly Ser Arg Ile Lys Leu Leu 115 120 125	384
40	CAC AGA ATT GGA TAT CCT TTA GCA AAG ATT CTT GCA TCC ATT GCA TAC His Arg Ile Gly Tyr Pro Leu Ala Lys Ile Leu Ala Ser Ile Ala Tyr 130 135 140	432
45	AAG AAG TCT TCA AGA TTG GTC GCA GAT CTT TCC TTT GGC AAA AAT GCT Lys Ser Ser Arg Leu Val Ala Asp Leu Ser Phe Gly Lys Asn Ala 145 150 155 160	480
50	GGT GAA CTT AAA GAG TGG GGA TGG AAA CAG GCA ATG GAT TAT ACA CCC Gly Glu Leu Lys Glu Trp Gly Trp Lys Gln Ala Met Asp Tyr Thr Pro 165 170 175	528
55	TCC TAC GTG GCA ATG TAC ACC TAC AGA ACT CTA ACG AAA GTG AAT CTT Ser Tyr Val Ala Met Tyr Thr Tyr Arg Thr Leu Thr Lys Val Asn Leu 180 185 190	576
60	GAA AAT ATC TTG GAG AAA ATA GAC TGT CCA ACA CTG ATT ATC GTT GGA Glu Asn Ile Leu Glu Lys Ile Asp Cys Pro Thr Leu Ile Ile Val Gly 195 200 205	624
65	GAA GAG GAT GCA CTA TTG CCC GTT AGC AAA TCA GTT GAG CTG AGC AGG Glu Glu Asp Ala Leu Leu Pro Val Ser Lys Ser Val Glu Leu Ser Arg 210 215 220	672
70	AGG ATA GAA AAC TCA AAG CTT GTG ATC ATC CCA AAC TCG GGG CAT TGC Arg Ile Glu Asn Ser Lys Leu Val Ile Ile Pro Asn Ser Gly His Cys 225 230 235 240	720
75	GTA ATG CTT GAG AGT CCA AGT GAG GTT AAT AGA GCA ATG GAC GAA TTC Val Met Leu Glu Ser Pro Ser Glu Val Asn Arg Ala Met Asp Glu Phe 245 250 255	768
80	ATT TCT TCA GCA CAG TTC TAA Ile Ser Ser Ala Gln Phe 260	774

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 756 NUCLEOTIDES
 (B) TYPE: NUCLEIC ACID
 5 (C) STRANDEDNESS: SINGLE
 (D) TOPOLOGY: LINEAR

(ii) MOLECULAR TYPE: GENOMIC DNA

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

	TTG AGA TTG AGG AAA TTT GAA GAG ATA AAC CTC GTT CTT TCG GGA GGA	48
	Leu Arg Leu Arg Lys Phe Glu Glu Ile Asn Leu Val Leu Ser Gly Gly	
15	1 5 10 15	
	GCT GCA AAG GGC ATA GCC CAC ATA GGT GTT TTG AAA GCT ATA AAC GAG	96
	Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys Ala Ile Asn Glu	
	20 25 30	
20	CTC GGT ATA AGG GTG AGG GCT TTA AGC GGG GTG AGC GCC CGG GCA ATC	144
	Leu Gly Ile Arg Val Arg Ala Leu Ser Gly Val Ser Ala Gly Ala Ile	
	35 40 45	
	GTT TCG GTC TTT TAT GCC TCA GGC TAC TCC CCT GAA GGG ATG TTC AGC	192
	Val Ser Val Phe Tyr Ala Ser Gly Tyr Ser Pro Glu Gly Met Phe Ser	
25	50 55 60	
	CTT CTG AAG AGG GTA AAC TGG CTG AAG CTG TTT AAG TTC AAG CCA CCT	240
	Leu Leu Lys Arg Val Asn Trp Leu Lys Leu Phe Lys Phe Lys Pro Pro	
	65 70 75 80	
30	CTG AAG GGA TTG ATA GGG TGG GAG AAG GCT ATA AGA TTC CTT GAG GAA	288
	Leu Lys Gly Leu Ile Gly Trp Glu Lys Ala Ile Arg Phe Leu Glu Glu	
	85 90 95	
	GTT CTC CCT TAC AGG AGA ATA GAA AAA CTT GAG ATA CCG ACG TAT ATA	336
	Val Leu Pro Tyr Arg Arg Ile Glu Lys Leu Glu Ile Pro Thr Tyr Ile	
	100 105 110	
35	TGC GCG ACG GAT TTA TAC TCG GGA AGG GCT CTA TAC CTC TCG GAA GGG	384
	Cys Ala Thr Asp Leu Tyr Ser Gly Arg Ala Leu Tyr Leu Ser Glu Gly	
	115 120 125	
40	AGT TTA ATC CCC GCA CTT CTC GGC AGC TGT GCA ATT CCC GGC ATA TTT	432
	Ser Leu Ile Pro Ala Leu Leu Gly Ser Cys Ala Ile Pro Gly Ile Phe	
	130 135 140	
	GAA CCC GTT GAG TAT AAG AAT TAC TTG CTC GTT GAC GGA GGT ATA GTT	480
	Glu Pro Val Glu Tyr Lys Asn Tyr Leu Leu Asp Gly Gly Ile Val	
	145 150 155 160	
45	AAC AAC CTT CCC GTT GAG CCC TTT CAG GAA AGC GGT ATT CCC ACC GTT	528
	Asn Asn Leu Pro Val Glu Pro Phe Gln Glu Ser Gly Ile Pro Thr Val	
	165 170 175	
50	TGC GTT GAT GTC CTT CCC ATA GAG CCG GAA AAG GAT ATA AAG AAC ATT	576
	Cys Val Asp Val Leu Pro Ile Glu Pro Glu Lys Asp Ile Lys Asn Ile	
	180 185 190	

55

EP 0 880 590 B1

	CTT CAC ATC CTT TTG AGG AGC TTC TTT CTT GCG GTC CGC TCA AAC TCC Leu His Ile Leu Leu Arg Ser Phe Phe Leu Ala Val Arg Ser Asn Ser 195 200 205	624
5	GAA AAG AGA AAG GAG TTT TGT GAC CTC GTT ATA GTT CCT GAG CTT GAG Glu Lys Arg Lys Glu Phe Cys Asp Leu Val Ile Val Pro Glu Leu Glu 210 215 220	672
	GAG TTC ACA CCC CTT GAT GTT AGA AAA GCG GAC CAA ATA ATG GAG AGG Glu Phe Thr Pro Leu Asp Val Arg Lys Ala Asp Gln Ile Met Glu Arg 225 230 235 240	720
10	GGA TAC ATA AAG GCC TTA GAG GTA CTT TCT GAA TAG Gly Tyr Ile Lys Ala Leu Glu Val Leu Ser Glu 245 250	768

15 (2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 894 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULAR TYPE: GENOMIC DNA

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

30	ATG TTT AAT ATC AAT GTC TTT GTT AAT ATA TCT TGG CTG TAT TTT TCA Met Phe Asn Ile Asn Val Phe Val Asn Ile Ser Trp Leu Tyr Phe Ser 1 5 10 15	48
	GGG ATA GTT ATG AAG ACT GTG GAA GAG TAT GCG CTA CTT GAA ACA GGC Gly Ile Val Met Lys Thr Val Glu Glu Tyr Ala Leu Leu Glu Thr Gly 20 25 30	96
35	GTA AGA GTG TTT TAT CGG TGT GTA ATC CCG GAG AAA GCT TTT AAC ACT Val Arg Val Phe Tyr Arg Cys Val Ile Pro Glu Lys Ala Phe Asn Thr 35 40 45	144
	TTG ATA ATA GGT TCA CAC GGA TTG GGG GCG CAC AGT GGA ATC TAC ATT Leu Ile Ile Gly Ser His Gly Leu Gly Ala His Ser Gly Ile Tyr Ile 50 55 60	192
40	AGT GTT GCT GAA GAA TTT GCT AGG CAC GGA TTT GGA TTC TGC ATG CAC Ser Val Ala Glu Glu Phe Ala Arg His Gly Phe Gly Phe Cys Met His 65 70 75 80	240
	GAT CAA AAG GGA CAT GGG AGA ACG GCA AGC GAT AGA GAA AGA GGG TAT Asp Gln Arg Gly His Gly Arg Thr Ala Ser Asp Arg Glu Arg Gly Tyr 85 90 95	288
	GTG GAG GGC TTT CAC AAC TTC ATA GAG GAT ATG AAG GCC TTC TCC GAT Val Glu Gly Phe His Asn Phe Ile Glu Asp Met Lys Ala Phe Ser Asp 100 105 110	336
50	TAT GCC AAG TGG CGC GTG GGA GGT GAC GAA ATA ATA TTG CTA GGA CAC Tyr Ala Lys Trp Arg Val Gly Gly Asp Glu Ile Ile Leu Leu Gly His 115 120 125	384
	AGT ATG GGC GGG CTG ATA GCG CTC TTA ACA GTT GCA ACT TAT AAA GAA Ser Met Gly Gly Leu Ile Ala Leu Leu Thr Val Ala Thr Tyr Lys Glu 130 135 140	432
55	ATC GCC AAG GGA GTT ATC GCG CTA GCC CCC GCC CTC CAA ATC CCC TTA	480

EP 0 880 590 B1

	Ile Ala Lys Gly Val Ile Ala Leu Ala Pro Ala Leu Gln Ile Pro Leu	
145	150 155 160	
5	ACC CCG GCT AGA AGA CTT GTT CTA AGC CTC GCG TCA AGG CTT GCC CCG Thr Pro Ala Arg Arg Leu Val Leu Ser Leu Ala Ser Arg Leu Ala Pro	528
	165 170 175	
	CAT TCT AAG ATC ACC TTA CAA AGG AGA TTG CCG CAG AAA CCA GAG GGT His Ser Lys Ile Thr Leu Gln Arg Arg Leu Pro Gln Lys Pro Glu Gly	576
10	180 185 190	
	TTT CAA AGA GCA AAA GAT ATA GAA TAC AGT CTG AGT GAA ATA TCA GTC Phe Gln Arg Ala Lys Asp Ile Glu Tyr Ser Leu Ser Glu Ile Ser Val	624
15	195 200 205	
	AAG CTC GTG GAC GAA ATG ATT AAA GCA TCA TCT ATG TTC TGG ACC ATA Lys Leu Val Asp Glu Met Ile Lys Ala Ser Ser Met Phe Trp Thr Ile	672
	210 215 220	
	GCA GGG GAA ATT AAT ACT CCC GTC CTG CTT ATT CAT GGG GAA AAA GAC Ala Gly Glu Ile Asn Thr Pro Val Leu Leu Ile His Gly Glu Lys Asp	720
20	225 230 235 240	
	AAT GTC ATA CCT CCG GAG GCG AGC AAA AAA GCC TAC CAA TTA ATA CCT Asn Val Ile Pro Pro Glu Ala Ser Lys Lys Ala Tyr Gln Leu Ile Pro	768
	245 250 255	
25	TCA TTC CCT AAA GAG TTG AAA ATA TAC CCC GAT CTT GGA CAC AAC TTG Ser Phe Pro Lys Glu Leu Lys Ile Tyr Pro Asp Leu Gly His Asn Leu	816
	260 265 270	
	TTT TTT GAA CCA GGC GCG GTG AAA ATC GTC ACA GAC ATT GTA GAG TGG Phe Phe Glu Pro Gly Ala Val Lys Ile Val Thr Asp Ile Val Glu Trp	864
	275 280 285	
30	GTG AAG AAT CTA CCC AGG GAA AAT CCT TAA Val Lys Asn Leu Pro Arg Glu Asn Pro	874
	290 295	

(2) INFORMATION FOR SEQ ID NO:28:

35 (i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 789 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

40 (ii) MOLECULE TYPE: GENOMIC DNA

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

	ATG GAG GTT TAC AAG GCC AAA TTC GGC GAA GCA AAG CTC GGC TGG GTC Met Glu Val Tyr Lys Ala Lys Phe Gly Glu Ala Lys Leu Gly Trp Val	48
	1 5 10 15	
50	GTT CTG GTT CAT GGC CTC GGC GAG CAC AGC GGA AGG TAT GGA AGA CTG Val Leu Val His Gly Leu Gly Glu His Ser Gly Arg Tyr Gly Arg Leu	96
	20 25 30	
	ATT AAG GAA CTC AAC TAT GCC GGC TTT GGA GTT TAC ACC TTC GAC TGG Ile Lys Glu Leu Asn Tyr Ala Gly Phe Gly Val Tyr Thr Phe Asp Trp	144
	35 40 45	
55	CCC GGC CAC GGG AAG AGC CCG GGC AAG AGA GGG CAC ACG AGC GTC GAG	192

	Pro Gly His Gly Lys Ser Pro Gly Lys Arg Gly His Thr Ser Val Glu	
	50 55 60	
5	GAG GCG ATG GAA ATC ATC GAC TCG ATA ATC GAG GAG ATC AGG GAG AAG Glu Ala Met Glu Ile Ile Asp Ser Ile Ile Glu Glu Ile Arg Glu Lys	240
	65 70 75 80	
	CCC TTC CTC TTC GGC CAC AGC CTC GGT GGT CTA ACT GTC ATC AGG TAC Pro Phe Leu Phe Gly His Ser Leu Gly Gly Leu Thr Val Ile Arg Tyr	288
10	85 90 95	
	GCT GAG ACG CGG CCC GAT AAA ATA CGG GGA TTA ATA GCT TCC TCG CCT Ala Glu Thr Arg Pro Asp Lys Ile Arg Gly Leu Ile Ala Ser Ser Pro	336
	100 105 110	
15	GCC CTC GCC AAG AGC CCG GAA ACG CCG GGC TTC ATG GTG GCC CTC GCG Ala Leu Ala Lys Ser Pro Glu Thr Pro Gly Phe Met Val Ala Leu Ala	384
	115 120 125	
	AAG TTC CTT GGA AAG ATC GCC CCG GGA GTT GTT CTC TCC AAC GGC ATA Lys Phe Leu Gly Lys Ile Ala Pro Gly Val Val Ser Asn Gly Ile	432
20	130 135 140	
	AAG CCG GAA CTC CTC TCG AGG AAC AGG GAC GCC GTG AGG AGG TAC GTT Lys Pro Glu Leu Leu Ser Arg Asn Arg Asp Ala Val Arg Arg Tyr Val	480
	145 150 155 160	
	GAA GAC CCA CTC GTC CAC GAC AGG ATT TCG GCC AAG CTG GGA AGG AGC Glu Asp Pro Leu Val His Asp Arg Ile Ser Ala Lys Leu Gly Arg Ser	528
25	165 170 175	
	ATC TTC GTG AAC ATG GAG CTG GCC CAC AGG GAG GCG GAC AAG ATA AAA Ile Phe Val Asn Met Glu Leu Ala His Arg Glu Ala Asp Ile Lys	576
	180 185 190	
30	GTC CCG ATC CTC CTT CTG ATC GGC ACT GGC GAT GTA ATA ACC CCG CCT Val Pro Ile Leu Leu Ile Gly Thr Gly Asp Val Ile Thr Pro Pro	624
	195 200 205	
	GAA GGC TCA CGC AGA CTC TTC GAG GAG CTG GCC GTC GAG AAC AAA ACC Glu Gly Ser Arg Arg Leu Phe Glu Glu Leu Ala Val Glu Asn Lys Thr	672
35	210 215 220	
	CTG AGG GAG TTC GAG GGG GCG TAC CAC GAG ATA TTT GAA GAC CCC GAG Leu Arg Glu Phe Glu Gly Ala Tyr His Glu Ile Phe Glu Asp Pro Glu	720
	225 230 235 240	
	TGG GCC GAG GAG TTC CAC GAA ACA ATT GTT AAG TGG CTG GTT GAA AAA Trp Ala Glu Glu Phe His Glu Thr Ile Val Lys Trp Leu Val Glu Lys	768
40	245 250 255	
	TCG TAC TCT TCG GCT CAA TAA Ser Tyr Ser Ser Ala Gln	775
	260	

(2) INFORMATION FOR SEQ ID NO:29:

45 (i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 750 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

50 (ii) MOLECULE TYPE: GENOMIC DNA

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

EP 0 880 590 B1

	TTG ATT GGC AAT TTG AAA TTG AAG AGG TTT GAA GAG GTT AAC TTA GTT Leu Ile Gly Asn Leu Lys Leu Lys Arg Phe Glu Glu Val Asn Leu Val 1 5 10 15	48
5	CTT TCG CGA GGG GCT GCC AAG GGT ATC GCC CAT ATA GGT GTT TTA AAA Leu Ser Gly Gly Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys 20 25 30	96
10	GCT CTG GAA GAG CTC GGT ATA AAG GTA AAG AGG CTC AGC GGG GTA AGT Ala Leu Glu Leu Gly Ile Lys Val Lys Arg Leu Ser Gly Val Ser 35 40 45	144
	GCT GGA CCT ATC GTT TCC GTC TTT TAC GCT TCG GGC TAC ACT CCC GAC Ala Gly Ala Ile Val Ser Val Phe Tyr Ala Ser Gly Tyr Thr Pro Asp 50 55 60	192
15	GAG ATG TTA AAA CTC CTG AAA GAG GTA AAC TGG CTC AAA CTT TTT AAG Glu Met Leu Lys Leu Lys Glu Val Asn Trp Leu Lys Leu Phe Lys 65 70 75 80	240
	TTC AAA ACA CCG AAA ATG GGC TTA ATG GGG TGG GAG AAG GCT GCA GAG Phe Lys Thr Pro Lys Met Gly Leu Met Gly Trp Glu Lys Ala Ala Glu 85 90 95	288
20	TTT TTG GAA AAA GAG CTC GGA GTT AAG AGG CTG GAA GAC CTG AAC ATA Phe Leu Glu Lys Glu Leu Gly Val Lys Arg Leu Glu Asp Leu Asn Ile 100 105 110	336
25	CCA ACC TAT CTT TGC TCG GCG GAT CTG TAC ACG GGA AAG GCT CTT TAC Pro Thr Tyr Leu Cys Ser Ala Asp Leu Tyr Thr Gly Lys Ala Leu Tyr 115 120 125	384
	TTC GGC AGA GGT GAC TTA ATT CCC GTG CTT CTC GGA AGT TGT TCC ATA Phe Gly Arg Gly Asp Leu Ile Pro Val Leu Leu Gly Ser Lys Ser Ile 130 135 140	432
30	CCC GGG ATT TTT GAA CCA GTT GAG TAC GAG AAT TTT CTA CTT GTT GAC Pro Gly Ile Phe Glu Pro Val Glu Tyr Glu Asn Phe Leu Leu Val Asp 145 150 155 160	480
	GGA GGT ATA GTG AAC AAC CTG CCC GTA GAA CCT TTG GAA AAG TTC AAA Gly Gly Ile Val Asn Asn Leu Pro Val Glu Pro Leu Glu Lys Phe Lys 165 170 175	528
35	GAA CCC ATA ATC GGG GTA GAT GTG CTT CCC ATA ACT CAA GAA AGA AAG Glu Pro Ile Ile Gly Val Asp Val Leu Pro Ile Thr Gln Glu Arg Lys 180 185 190	576
40	ATT AAA AAT ATA CTC CAC ATC CTT ATA AGG AGC TTC TTT CTG GCG GTT Ile Lys Asn Ile Leu His Ile Leu Ile Arg Ser Phe Phe Leu Ala Val 195 200 205	624
	CGT TCC AAT TCG GAA AAG AGA AAG GAG TTC TGC AAC GTA GTT ATA GAA Arg Ser Asn Ser Glu Lys Arg Lys Glu Phe Phe Cys Asn Val Val Ile Glu 210 215 220	672
45	CCT CCC CTT GAA GAG TTC TCT CCT CTG GAC GTA AAT AAG GCG GAC GAG Pro Pro Leu Glu Phe Ser Pro Leu Asp Val Asn Lys Ala Asp Glu 225 230 235 240	720
50	ATA TTC TGC GGG GAT ATG AGA GCA CTT TAA Ile Phe Cys Gly Asp Met Arg Ala Leu 245	730

55 (2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS

EP 0 880 590 B1

- (A) LENGTH: 1017 NUCLEOTIDES
 (B) TYPE: NUCLEIC ACID
 (C) STRANDEDNESS: SINGLE
 (D) TOPOLOGY: LINEAR

5

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ IN NO:30:

10

ATG CCA GCT AAT GAC TCA CCC ACG ATC GAC TTT AAT CCT CGC GGC ATT	48
Met Pro Ala Asn Asp Ser Pro Thr Ile Asp Phe Asn Pro Arg Gly Ile	
1 5 10 15	

15

CCT CGC AAC GCT CAC GCA CAG GTT ATT TTA GCG ACT TCC GGC TTG CGC	96
Leu Arg Asn Ala His Ala Gln Val Ile Leu Ala Thr Ser Gly Leu Arg	
20 25 30	

AAA GCG TTT TTG AAA CGC ACG CAC AAG AGC TAC CTC AGC ACT GCC CAA	144
Lys Ala Phe Leu Lys Arg Thr His Lys Ser Tyr Leu Ser Thr Ala Gln	
35 40 45	

20

TGG CTG GAG CTC GAT GCC GGC AAC GGA GTT ACC TTG GCC GGA GAG CTT	192
Trp Leu Glu Leu Asp Ala Gly Asn Gly Val Thr Leu Ala Gly Glu Leu	
50 55 60	

25

AAC ACA GCG CCT GCA ACT GCA TCC TCC TCC CAC CGG CAC AAG AAC 240	
Asn Thr Ala Pro Ala Thr Ala Ser Ser Ser His Pro Ala His Lys Asn	
65 70 75 80	

ACT CTG GTT ATT GTG CTG CAC GGC TGG GAA GGC TCC AGC CAG TCG GCC	288
Thr Leu Val Ile Val Leu His Gly Trp Glu Gly Ser Ser Gln Ser Ala	
85 90 95	

30

TAT GCG ACC TCC GCT GGC AGC ACG CTT TTC GAC AAT GGG TTC GAC ACT	336
Tyr Ala Thr Ser Ala Gly Ser Thr Leu Phe Asp Asn Gly Phe Asp Thr	
100 105 110	

TTT CGC CTT AAT TTT CGC GAT CAC GGC GAC ACC TAC CAC TTA AAC CGC	384
Phe Arg Leu Asn Phe Arg Asp His Gly Asp Thr Tyr His Leu Asn Arg	
115 120 125	

35

GGC ATA TTT AAC TCA TCG CTG ATT GAC GAA GTA GTG GGC GCA GTC AAA	432
Gly Ile Phe Asn Ser Ser Leu Ile Asp Glu Val Val Gly Ala Val Lys	
130 135 140	

40

GCC ATC CAG CAG CAA ACC GAC TAC GAC AAG TAT TGC CTG ATG GGG TTC	480
Ala Ile Gln Gln Thr Asp Tyr Asp Lys Tyr Cys Leu Met Gly Phe	
145 150 155 160	

TCA CTG GGT GGG AAC TTT GCC TTG CGC GTC GCG GTG CGG GAA CAG CAT	528
Ser Leu Gly Gly Asn Phe Ala Leu Arg Val Ala Val Arg Glu Gln His	
165 170 175	

45

CTC GCT AAA CCG CTA GCG GGC GTG CTC GCC GTA TGC CCG GTA CTC GAC	576
Leu Ala Lys Pro Leu Ala Gly Val Leu Ala Val Cys Pro Val Leu Asp	
180 185 190	

CCC GCA CAC ACC ATG ATG GCC CTA AAC CGA GGT GCG TTT TTC TAC GGC	624
---	-----

50

55

EP 0 880 590 B1

	Pro Ala His Thr Met Met Ala Leu Asn Arg Gly Ala Phe Phe Tyr Gly		
	195 200 205		
5	CGC TAT TTT GCG CAT AAA TGG AAG CGC TCG TTA ACC GCA AAA CTT GCA		672
	Arg Tyr Phe Ala His Lys Trp Lys Arg Ser Leu Thr Ala Lys Leu Ala		
	210 215 220 225		
	GCT TTC CCA GAC TAC AAA TAC GGC AAA GAT TTA AAA TCG ATA CAC ACG		720
	Ala Phe Pro Asp Tyr Lys Tyr Gly Lys Asp Leu Lys Ser Ile His Thr		
10	230 235 240		
	CTT GAT GAG TTA AAC AAC TAT TTC ATT CCC CGC TAC ACC GGC TTC AAC		768
	Leu Asp Glu Leu Asn Asn Tyr Phe Ile Pro Arg Tyr Thr Gly Phe Asn		
	245 250 255		
15	TCA GTC TCC GAA TAC TTC AAA AGT TAC ACG CTC ACC GGG CAG AAG CTC		816
	Ser Val Ser Glu Tyr Phe Lys Ser Tyr Thr Leu Thr Gly Gln Lys Leu		
	260 265 270		
	GCG TTT CTC AAC TGC CCC AGT TAC ATT CTG GCA GCT GGC GAC GAC CCA		864
	Ala Phe Leu Asn Cys Pro Ser Tyr Ile Leu Ala Ala Gly Asp Asp Pro		
	275 280 285		
20	ATA ATT CCA GCA TCC GAC TTT CAG AAA ATA GCC AAG CCT GCG AAT CTG		912
	Ile Ile Pro Ala Ser Asp Phe Gln Lys Ile Ala Lys Pro Ala Asn Leu		
	290 295 300 305		
	CAC ATA ACA GTA ACG CAA CAA GGT TCT CAT TGC GCA TAC CTG GAA AAC		960
	His Ile Thr Val Thr Gln Gln Gly Ser His Cys Ala Tyr Leu Glu Asn		
25	310 315 320		
	CTG CAT AAA CCT AGT GCT GCC GAC AAA TAT GCG GTG AAA TTA TTT GGA		1,008
	Leu His Lys Pro Ser Ala Ala Asp Lys Tyr Ala Val Lys Leu Phe Gly		
	325 330 335		
30	GCC TGT TGA		1,111
	Ala Cys		

(2) INFORMATION FOR SEQ ID NO:31:

35 (i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 936 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- 40 (D) TOPOLOGY: LINEAR

45 (ii) MOLECULE TYPE: GENOMIC DNA .

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

50

55

EP 0 880 590 B1

ATG CTT GAT ATG CCA ATC GAC CCT GTT TAC TAC CAG CTT GCT GAG TAT Met Leu Asp Met Pro Ile Asp Pro Val Tyr Tyr Gln Leu Ala Glu Tyr 1 5 10 15	48
5 TTC GAC AGT CTG CCG AAG TTC GAC CAG TTT TCC TCG GCC AGA GAG TAC Phe Asp Ser Leu Pro Lys Phe Asp Gln Phe Ser Ser Ala Arg Glu Tyr 20 25 30	96
10 AGG GAG GCG ATA AAT CGA ATA TAC GAG GAG AGA AAC CGG CAG CTG AGC Arg Glu Ala Ile Asn Arg Ile Tyr Glu Glu Arg Asn Arg Gln Leu Ser 35 40 45	144
CAG CAT GAG AGG GTT GAA AGA GTT GAG GAC AGG ACG ATT AAG GGG AGG Gln His Glu Arg Val Glu Arg Val Glu Asp Arg Thr Ile Lys Gly Arg 50 55 60	192

15

20

25

30

35

40

45

50

55

EP 0 880 590 B1

	AAC GGA GAC ATC AGA GTC AGA GTT TAC CAG CAG AAG CCC GAT TCC CCG Asn Gly Asp Ile Arg Val Arg Val Tyr Gln Gln Lys Pro Asp Ser Pro 65 70 75 80	240
5	GGT CTG GTT TAC TAT CAC GGT GGT GGA TTT GTG ATT TGC AGC ATC GAG Val Leu Val Tyr Tyr His Gly Gly Gly Phe Val Ile Cys Ser Ile Glu 85 90 95	288
10	TCG CAC GAC GCC TTA TGC AGG AGA AYY GCG AGA CTT TCA AAC TCT ACC Ser His Asp Ala Leu Cys Arg Arg Ile Ala Arg Leu Ser Asn Ser Thr 100 105 110	336
	GTA GTC TCC GTG GAT TAC AGG CTC GCT CCT GAG CAC AAG TTT CCC CCC Val Val Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Ala 115 120 125	384
15	CCA GTT TAT CAT TGC TAC GAT GCG ACC AAG TGG GTT GCT GAG AAC CGG Ala Val Tyr Asp Cys Tyr Asp Ala Thr Lys Trp Val Ala Glu Asn Ala 130 135 140	432
20	GAG GAG CTG AGG ATT GAC CCG TCA AAA ATC TTC GTT GGG GGG GAC AGT Glu Glu Leu Arg Ile Asp Pro Ser Lys Ile Phe Val Gly Gly Asp Ser 145 150 155 160	480
	GCG GGA CGG AAT CTT GCC CCG GCG CTT TCA ATA ATG GCG AGA GAC AGC Ala Gly Gly Asn Leu Ala Ala Val Ser Ile Met Ala Arg Asp Ser 165 170 175	528
25	GGA GAA GAT TTC ATA AAG CAT CAA ATT CTA ACT TAC CCC GTT GTG AAC Gly Glu Asp Phe Ile Lys His Gln Ile Leu Ile Tyr Pro Val Val Asn 180 185 190	576
	TTT GTA GCC CCC ACA CCA TCG CTT CTG GAG TTT GGA GAG GGG CTG TCG Phe Val Ala Pro Thr Pro Ser Leu Leu Glu Phe Gly Glu Gly Leu Trp 195 200 205	624
30	ATT CTC GAC CAG AAG ATA ATG AGT TGG TTC TCG GAG CAG TAC TTC TCC Ile Leu Asp Gln Lys Ile Met Ser Trp Phe Ser Glu Gln Tyr Phe Ser 210 215 230	672
35	AGA GAG GAA GAT AAG TTC AAG CCC CTC GCC TCC GTC ATC ITT GCG GAC Arg Glu Glu Asn Lys Phe Asn Pro Leu Ala Ser Val Ile Phe Ala Asp 235 240 245 250	720
	CTT GAG AAC CTA CCT CCT GCG CTG ATC ATA ACC GCC GAA TAC GAC CCG Leu Glu Asn Leu Pro Pro Ala Leu Ile Ile Thr Ala Glu Tyr Asp Pro 255 260 265	768
40	CTG AGA GAT GAA GGA GAA GTT TTC GGG CAG ATG CTG AGA AGA GCA GGT Leu Arg Asp Glu Gly Glu Val Phe Gly Glu Met Leu Arg Arg Ala Gly 270 275 280	816
	GTT GAG GCG AGC ATC GTC AGA TAC AGA GGC GTG CTT CAC GGA TTC ATC Val Glu Ala Ser Ile Val Arg Tyr Arg Gly Val Leu His Gly Phe Ile 285 290 295	864
45	AAT TAC TAT CCC GTG CTG AAG GCT GCG AGG GAT GCG ATA AAC CAG ATT Asn Tyr Tyr Pro Val Leu Lys Ala Ala Arg Asp Ala Ile Asn Gln Ile 300 305 310	912
50	GCC GCT CTT CTT GTG TTC GAC TAG Ala Ala Leu Leu Val Phe Asp 315 320	936

(2) INFORMATION FOR SEQ ID NO:32:

55 (i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 918 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID

(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

	ATG CCC CTA GAT CCT AGA ATT AAA AAG TTA CTA GAA TCA GCT CTT ACT	48
10	Met Pro Leu Asp Pro Arg Ile Lys Lys Leu Leu Glu Ser Ala Leu Thr	
	5 10 15	
	ATA CCA ATT GGT AAA GCC CCA GTC GAA GAG GTC AGA AAG ATA TTT AGG	96
	Ile Pro Ile Gly Lys Ala Pro Val Glu Glu Val Arg Lys Ile Phe Arg	
	20 25 30	
15	CAA TTA GCG TCG GCA GCT CCC AAA GTC GAA GTT GGA AAA GTC GAA GAT	144
	Gln Leu Ala Ser Ala Ala Pro Lys Val Glu Val Gly Lys Val Glu Asp	
	35 40 45	
20	ATA AAA ATA CCA GGC AGT GAA ACC GTT ATA AAC GCT AGA GTG TAT TTT	192
	Ile Lys Ile Pro Gly Ser Glu Thr Val Ile Asn Ala Arg Val Tyr Phe	
	50 55 60	
	CCG AAG AGT AGC GGT CCT TAT GGT GTT CTA GTG TAT CTT CAT GGA GGC	240
	Pro Lys Ser Ser Gly Pro Tyr Gly Val Leu Val Tyr Leu His Gly Gly	
	65 70 75 80	
25	GGT TTT GTA ATA GGC GAT GTG GAA TCT TAT GAC CCA TTA TGT AGA GCA	288
	Gly Phe Val Ile Gly Asp Val Glu Ser Tyr Asp Pro Leu Cys Arg Ala	
	85 90 95	
	ATT ACA AAT GCG TGC AAT TGC GTT GTA TCA GTG GAC TAT AGG TTA	336
	Ile Thr Asn Ala Cys Asn Cys Val Val Val Ser Val Asp Tyr Arg Leu	
	100 105 110	
30	GCT CCA GAA TAC AAG TTT CCT TCT GCA GTT ATC GAT TCA TTT GAC GCT	384
	Ala Pro Glu Tyr Lys Phe Pro Ser Ala Val Ile Asp Ser Phe Asp Ala	
	115 120 125	
35	ACT AAT TGG GTT TAT AAC AAT TTA GAT AAA TTT GAT GGA AAG ATG GGA	432
	Thr Asn Tyr Val Tyr Asn Asn Leu Asp Lys Phe Asp Gly Lys Met Gly	
	130 135 140	
	GTT GCG ATT GCG GGA GAT AGT GCT GGA GGA AAT TTG GCA GCG GTT GTA	480
	Val Ala Ile Ala Gly Asp Ser Ale Gly Gly Asn Leu Ala Ala Val Val	
	145 150 155 160	
40	GCT CCT CTT TCA AAG GGT AAA ATT AAT TTG AAG TAT CAA ATA CTG GTT	528
	Ala Leu Leu Ser Lys Gly Lys Ile Asn Leu Lys Tyr Gln Ile Leu Val	
	165 170 175	
	TAC CCA GCG GTA AGT TTA GAT AAC GTT TCA AGA TCC ATG ATA GAG TAC	576
	Tyr Pro Ala Val Ser Leu Asp Asn Val Ser Arg Ser Met Ile Glu Tyr	
	180 185 190	
45	TCT GAT GGG TTC TTC CTT ACC AGA GAG CAT ATA GAG TGG TTC GGT TCT	624
	Ser Asp Gly Phe Phe Leu Thr Arg Glu His Ile Glu Trp Phe Gly Ser	
	195 200 205	

50

55

EP 0 880 590 B1

CAA TAC TTA CGA AGC CCT GCA GAT TTG CTA GAC TTT AGG TTC TCT CCA Gln Tyr Leu Arg Ser Pro Ala Asp Leu Leu Asp Phe Arg Phe Ser Pro 210 215 220	672
5 ATT CTG GCG CAA GAT TTC AAC GGA TTA CCT CCA GCC TTG ATA ATA ACA Ile Leu Ala Gln Asp Phe Asn Gly Leu Pro Pro Ala Leu Ile Ile Thr 225 230 235 240	720
10 GCA GAA TAC GAT CCA CTA AGG GAT CAA GGA GAA GCG TAT GCA AAT AAA Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Ala Tyr Ala Asn Lys 245 250 255	768
15 CTA CTA CAA GCT GGA GTC TCA GTT ACT AGT GTG AGA TTT AAC AAC GTT Leu Leu Gln Ala Gly Val Ser Val Thr Ser Val Arg Phe Asn Asn Val 260 265 270	816
ATA CAC GGA TTC CTC TCA TTC TTT CCG TTG ATG GAG CAA GGA AGA GAT Ile His Gly Phe Leu Ser Phe Phe Pro Leu Met Glu Gln Gly Arg Asp 275 280 285	864
20 GCT ATA GGT CTG ATA GGG TCT GTG TTA AGA CGA GTA TTT TAT GAT AAA Ala Ile Gly Leu Ile Gly Ser Val Leu Arg Arg Val Phe Tyr Asp Lys 290 295 300	912-
ATT TAA Ile 305	918

25 (2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 184 AMINO ACIDS
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

40

45

50

55

Met Ser Leu Asn Lys His Ser Trp Met Asp Met Ile Ile Phe Ile Leu
 1 5 10 15

5 Ser Phe Ser Phe Pro Leu Thr Met Ile Ala Leu Ala Ile Ser Met Ser
 20 25 30

Ser Trp Phe Asn Ile Trp Asn Asn Ala Leu Ser Asp Leu Gly His Ala
 35 40 45

Val Lys Ser Ser Val Ala Pro Ile Phe Asn Leu Gly Leu Ala Ile Gly
 10 50 55 60

Gly Ile Leu Ile Val Ile Val Gly Leu Arg Asn Leu Tyr Ser Trp Ser
 65 70 75 80

Arg Val Lys Gly Ser Leu Ile Ile Ser Met Gly Val Phe Leu Asn Leu
 15 85 90 95

Ile Gly Val Phe Asp Glu Val Tyr Gly Trp Ile His Phe Leu Val Ser
 100 105 110

Val Leu Phe Phe Leu Ser Ile Ile Ala Tyr Phe Ile Ala Ile Ser Ile
 20 115 120 125

Leu Asp Lys Ser Trp Ile Ala Val Leu Leu Ile Ile Gly His Ile Ala
 130 135 140

25 Met Trp Tyr Leu His Phe Ala Ser Glu Ile Pro Arg Gly Ala Ala Ile
 145 150 155 160

30 Pro Glu Leu Leu Ala Val Phe Ser Phe Leu Pro Phe Tyr Ile Arg Asp
 165 170 175

Tyr Phe Lys Ser Tyr Thr Lys Arg
 180

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 346 AMINO ACIDS
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

50

55

Met Lys Leu Leu Glu Pro Thr Asn Thr Ser Tyr Thr Leu Leu Gln Asp
 1 5 10 15

5 Leu Ala Leu His Phe Ala Phe Tyr Trp Phe Leu Ala Val TYr Thr Trp
 20 25 30

Leu Pro Gly Val Leu Val Arg Gly Val Ala Val Asp Thr Gly Val Ala
 35 40 45

10 Arg Val Pro Gly Leu Gly Arg Arg Gly Lys Arg Leu Leu Leu Ala Ala
 50 55 60

Val Ala Val Leu Ala Leu Val Val Ser Val Val Val Pro Ala Tyr Val
 65 70 75 80

15 Ala Tyr Ser Ser Leu His Pro Glu Ser Cys Arg Pro Val Ala Pro Glu
 85 90 95

Gly Leu Thr Tyr Lys Glu Phe Ser Val Thr Ala Glu Asp Gly Leu Val
 100 105 110

20 Val Arg Gly Trp Cal Leu Gly Pro Gly Ala Gly Gly Asn Pro Val Phe
 115 120 125

Val Leu Met His Gly Tyr Thr Gly Cys Arg Ser Ala Pro Tyr Met Ala
 130 135 140

25 Val Leu Ala Arg Glu Leu Val Glu Trp Gly Tyr Pro Val Val Val Phe
 145 150 155 160

Asp Phe Arg Gly His Gly Glu Ser Gly Gly Ser Thr Thr Ile Gly Pro
 165 170 175

30 Arg Glu Val Leu Asp Ala Arg Ala Val Val Gly Tyr Val Ser Glu Arg
 180 185 190

Phe Pro Gly Arg Arg Ile Ile Leu Val Gly Phe Ser Met Gly Gly Ala
 195 200 205

35 Val Ala Ile Val Glu Gly Ala Gly Asp Pro Arg Val Tyr Ala Val Ala
 210 215 220

Ala Asp Ser Pro Tyr Tyr Arg Leu Arg Asp Val Ile Pro Arg Trp Leu
 225 230 235 240

40 Glu Tyr Lys Thr Pro Leu Pro Gly Trp Val Gly Val Leu Ala Gly Phe
 245 250 255

Tyr Gly Arg Leu Met Ala Gly Val Asp Leu Gly Phe Gly Pro Ala Gly
 260 265 270

45 Val Glu Arg Val Asp Lys Pro Leu Leu Val Val Tyr Gly Pro Arg Asp
 275 280 285

Pro Leu Val Thr Arg Asp Glu Ala Arg Ser Leu Ala Ser Arg Ser Pro
 290 295 300

50 Cys Gly Arg Leu Val Glu Val Pro Gly Ala Gly His Val Glu Ala Val
 305 310 315 320

Asp Val Leu Gly Pro Gly Arg Tyr Ala Asp Met Leu Ile Glu Leu Ala
 325 330 335

55 His Glu Glu Cys Pro Pro Gly Ala Gly Gly
 340 345

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 262 AMINO ACIDS
 - (B) TYPE: AMINO ACID
 - (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Met Pro Tyr Val Arg Asn Gly Gly Val Asn Ile Tyr Tyr Glu Leu Val
 1 5 10 15
 Asp Gly Pro Glu Pro Pro Ile Val Phe Val His Gly Trp Thr Ala Asn
 20 25 30
 Met Asn Phe Trp Lys Glu Gln Arg Arg Tyr Phe Ala Gly Arg Asn Met
 35 40 45
 Met Leu Phe Val Asp Asn Arg Gly His Gly Arg Ser Asp Lys Pro Leu
 50 55 60
 Gly Tyr Asp Phe Tyr Arg Phe Glu Asn Phe Ile Ser Asp Leu Asp Ala
 65 70 75 80
 Val Val Arg Glu Thr Gly Val Glu Lys Phe Val Leu Val Gly His Ser
 85 90 95
 Phe Gly Thr Met Ile Ser Met Lys Tyr Cys Ser Glu Tyr Arg Asn Arg
 100 105 110
 Val Leu Ala Leu Ile Leu Ile Gly Gly Ser Arg Ile Lys Leu Leu
 115 120 125
 His Arg Ile Gly Tyr Pro Leu Ala Lys Ile Leu Ala Ser Ile Ala Tyr
 130 135 140
 Lys Lys Ser Ser Arg Leu Val Ala Asp Leu Ser Phe Gly Lys Asn Ala
 145 150 155 160
 Gly Glu Leu Lys Glu Trp Gly Trp Lys Gln Ala Met Asp Tyr Thr Pro
 165 170 175

 Ser Tyr Val Ala Met Tyr Thr Tyr Arg Thr Leu Thr Lys Val Asn Leu
 180 185 190
 Glu Asn Ile Leu Glu Lys Ile Asp Cys Pro Thr Leu Ile Ile Val Gly
 195 200 205
 Glu Glu Asp Ala Leu Leu Pro Val Ser Lys Ser Val Glu Leu Ser Arg
 210 215 220
 Arg Ile Glu Asn Ser Lys Leu Val Ile Ile Pro Asn Ser Gly His Cys
 225 230 235 240
 Val Met Leu Glu Ser Pro Ser Glu Val Asn Arg Ala Met Asp Glu Phe
 245 250 255
 Ile Ser Ser Ala Gln Phe
 260

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS

- 5
 (A) LENGTH: 251 AMINO ACIDS
 (B) TYPE: AMINO ACID
 (D) TOPOLOGY: LINEAR

10 (ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

15 Leu Arg Leu Arg Lys Phe Glu Glu Ile Asn Leu Val Leu Ser Gly Gly
 1 5 10 15
 Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys Ala Ile Asn Glu
 20 25 30
 20 Leu Gly Ile Arg Val Arg Ala Leu Ser Gly Val Ser Ala Gly Ala Ile
 35 40 45
 25 Val Ser Val Phe Tyr Ala Ser Gly Tyr Ser Pro Glu Gly Met Phe Ser
 50 55 60
 30 Leu Leu Lys Arg Val Asn Trp Leu Lys Leu Phe Lys Phe Lys Pro Pro
 65 70 75 80
 35 Leu Lys GIy Leu Ile Gly Trp Glu Lys Ala Ile Arg Phe Leu Glu Glu
 85 90 95
 40 Val Leu Pro Tyr Arg Arg Ile Glu Lys Leu Glu Ile Pro Thr Tyr Ile
 100 105 110
 45 Cys Ala Thr Asp Leu Tyr Ser Gly Arg Ala Leu Tyr Leu Ser Glu Gly
 115 120 125
 50 Ser Leu Ile Pro Ala Leu Leu Gly Ser Cys Ala Ile Pro Gly Ile Phe
 130 135 140
 55 Glu Pro Val Glu Tyr Lys Asn Tyr Leu Leu Val Asp Gly Gly Ile Val
 145 150 155 160
 60 Asn Asn Leu Pro Val Glu Pro Phe Gln Glu Ser Gly Ile Pro Thr Val
 165 170 175
 65 Cys Val Asp Val Leu Pro Ile Glu Pro Glu Lys Asp Ile Lys Asn Ile
 180 185 190
 70 Leu His Ile Leu Leu Arg Ser Phe Phe Leu Ala Val Arg Ser Asn Ser
 195 200 205
 75 Glu Lys Arg Lys Glu Phe Cys Asp Leu Val Ile Val Pro Glu Leu Glu
 210 215 220
 80 Glu Phe Thr Pro Leu Asp Val Arg Lys Ala Asp Gln Ile Met Glu Arg
 225 230 235 240
 85 Gly Tyr Ile Lys Ala Leu Glu Val Leu Ser Glu
 245 250

55

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 297 AMINO ACIDS
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

5 (ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

10	Met Phe Asn Ile Asn Val Phe Val Asn Ile Ser Trp Leu Tyr Phe Ser	1	5	10	15
	Gly Ile Val Met Lys Thr Val Glu Glu Tyr Ala Leu Leu Glu Thr Gly	20	25	30	
15	Val Arg Val Phe Tyr Arg Cys Val Ile Pro Glu Lys Ala Phe Asn Thr	35	40	45	
	Leu Ile Ile Gly Ser His Gly Leu Gly Ala His Ser Gly Ile Tyr Ile	50	55	60	
20	Ser Val Ala Glu Glu Phe Ala Arg His Gly Phe Gly Phe Cys Met His	65	70	75	80
	Asp Gln Arg Gly His Gly Arg Thr Ala Ser Asp Arg Glu Arg Gly Tyr	85	90	95	
25	Val Glu Gly Phe His Asn Phe Ile Glu Asp Met Lys Ala Phe Ser Asp	100	105	110	
	Tyr Ala Lys Trp Arg Val Gly Gly Asp Glu Ile Ile Leu Leu Gly His	115	120	125	
30	Ser Met Gly Gly Leu Ile Ala Leu Leu Thr Val Ala Thr Tyr Lys Glu	130	135	140	
	Ile Ala Lys Gly Val Ile Ala Leu Ala Pro Ala Leu Gln Ile Pro Leu	145	150	155	160
35	Thr Pro Ala Arg Arg Leu Val Leu Ser Leu Ala Ser Arg Leu Ala Pro	165	170	175	
	His Ser Lys Ile Thr Leu Gln Arg Arg Leu Pro Gln Lys Pro Glu Gly	180	185	190	
40	Phe Gln Arg Ala Lys Asp Ile Glu Tyr Ser Leu Ser Glu Ile Ser Val	195	200	205	
	Lys Leu Val Asp Glu Met Ile Lys Ala Ser Ser Met Phe Trp Thr Ile	210	215	220	
45					
	Ala Gly Glu Ile Asn Thr Pro Val Leu Leu Ile His Gly Glu Lys Asp	225	230	235	240
50	Asn Val Ile Pro Pro Glu Ala Ser Lys Lys Ala Tyr Gln Leu Ile Pro	245	250	255	
	Ser Phe Pro Lys Glu Leu Lys Ile Tyr Pro Asp Leu Gly His Asn Leu	260	265	270	
55	Phe Phe Glu Pro Gly Ala Val Lys Ile Val Thr Asp Ile Val Glu Trp	275	280	285	
	Val Lys Asn Leu Pro Arg Glu Asn Pro	290	295		

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 262 AMINO ACIDS
 - (B) TYPE: AMINO ACID
 - (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Met Glu Val Tyr Lys Ala Lys Phe Gly Glu Ala Lys Leu Gly Trp Val
 1 5 10 15

Val Leu Val His Gly Leu Gly Glu His Ser Gly Arg Tyr Gly Arg Leu
 20 25 30

Ile Lys Glu Leu Asn Tyr Ala Gly Phe Gly Val Tyr Thr Phe Asp Trp
 35 40 45

Pro Gly His Gly Lys Ser Pro Gly Lys Arg Gly His Thr Ser Val Glu
 50 55 60

Glu Ala Met Glu Ile Ile Asp Ser Ile Ile Glu Glu Ile Arg Glu Lys
 65 70 75 80

Pro Phe Leu Phe Gly His Ser Leu Gly Gly Leu Thr Val Ile Arg Tyr
 85 90 95

Ala Glu Thr Arg Pro Asp Lys Ile Arg Gly Leu Ile Ala Ser Ser Pro
 100 105 110

Ala Leu Ala Lys Ser Pro Glu Thr Pro Gly Phe Met Val Ala Leu Ala
 115 120 125

Lys Phe Leu Gly Lys Ile Ala Pro Gly Val Val Leu Ser Asn Gly Ile
 130 135 140

Lys Pro Glu Leu Leu Ser Arg Asn Arg Asp Ala Val Arg Arg Tyr Val
 145 150 155 160

Glu Asp Pro Leu Val His Asp Arg Ile Ser Ala Lys Leu Gly Arg Ser
 165 170 175

Ile Phe Val Asn Met Glu Leu Ala His Arg Glu Ala Asp Lys Ile Lys
 180 185 190

Val Pro Ile Leu Leu Ile Gly Thr Gly Asp Val Ile Thr Pro Pro
 195 200 205

Glu Gly Ser Arg Arg Leu Phe Glu Glu Leu Ala Val Glu Asn Lys
 210 215 220

Leu Arg Glu Phe Glu Gly Ala Tyr His Glu Ile Phe Glu Asp Pro
 225 230 235

Trp Ala Glu Glu Phe His Glu Thr Ile Val Lys Trp Leu Val Glu
 245 250 255

Ser Tyr Ser Ser Ala Gln
 260

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS

- 5
 (A) LENGTH: 249 AMINO ACIDS
 (B) TYPE: AMINO ACID
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

10
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Leu Ile Gly Asn Leu Lys Leu Lys Arg Phe Glu Glu Val Asn Leu Val
 1 5 10 15

15 Leu Ser Gly Gly Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys
 20 25 30

Ala Leu Glu Glu Leu Gly Ile Lys Val Lys Arg Leu Ser Gly Val Ser
 35 40 45

20 Ala Gly Ala Ile Val Ser Val Phe Tyr Ala Ser Gly Tyr Thr Pro Asp
 50 55 60

Glu Met Leu Lys Leu Leu Lys Glu Val Asn Trp Leu Lys Leu Phe Lys
 -65 70 75 80

25 Phe Lys Thr Pro Lys Met Gly Leu Met Gly Trp Glu Lys Ala Ala Glu
 85 90 95

Phe Leu Glu Lys Glu Leu Gly Val Lys Arg Leu Glu Asp Leu Asn Ile
 100 105 110

30 Pro Thr Tyr Leu Cys Ser Ala Asp Leu Tyr Thr Gly Lys Ala Leu Tyr
 115 120 125

Phe Gly Arg Gly Asp Leu Ile Pro Val Leu Leu Gly Ser Lys Ser Ile
 130 135 140

35 Pro Gly Ile Phe Glu Pro Val Glu Tyr Glu Asn Phe Leu Leu Val Asp
 145 150 155 160

Gly Gly Ile Val Asn Asn Leu Pro Val Glu Pro Leu Glu Lys Phe Lys
 165 170 175

40 Glu Pro Ile Ile Gly Val Asp Val Leu Pro Ile Thr Gln Glu Arg Lys
 180 185 190

Ile Lys Asn Ile Leu His Ile Leu Ile Arg Ser Phe Phe Leu Ala Val
 195 200 205

45 Arg Ser Asn Ser Glu Lys Arg Lys Glu Phe Cys Asn Val Val Ile Glu
 210 215 220

50 Pro Pro Leu Glu Glu Phe Ser Pro Leu Asp Val Asn Lys Ala Asp Glu
 225 230 235 240

Ile Phe Cys Gly Asp Met Arg Ala Leu
 245

55
 (2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 339 AMINO ACIDS
 (B) TYPE: AMINO ACID
 (D) TOPOLOGY: LINEAR

5 (ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

10	Met Pro Ala Asn Asp Ser Pro Thr Ile Asp Phe Asn Pro Arg Gly Ile		
	1 5 10 15		
	Leu Arg Asn Ala His Ala Gln Val Ile Leu Ala Thr Ser Gly Leu Arg		
	20 25 30		
15	Lys Ala Phe Leu Lys Arg Thr His Lys Ser Tyr Leu Ser Thr Ala Gln		
	35 40 45		
	Trp Leu Glu Leu Asp Ala Gly Asn Gly Val Thr Leu Ala Gly Glu Leu		
	50 55 60		
20	Asn Thr Ala Pro Ala Thr Ala Ser Ser His Pro Ala His Lys Asn		
	65 70 75 80		
	Thr Leu Val Ile Val Leu His Gly Trp Glu Gly Ser Ser Gln Ser Ala		
	85 90 95		
25	Tyr Ala Thr Ser Ala Gly Ser Thr Leu Phe Asp Asn Gly Phe Asp Thr		
	100 105 110		
	Phe Arg Leu Asn Phe Arg Asp His Gly Asp Thr Tyr His Leu Asn Arg		
	115 120 125		
30	Gly Ile Phe Asn Ser Ser Leu Ile Asp Glu Val Val Gly Ala Val Lys		
	130 135 140		
	Ala Ile Gln Gln Gln Thr Asp Tyr Asp Lys Tyr Cys Leu Met Gly Phe		
	145 150 155 160		
35	Ser Leu Gly Gly Asn Phe Ala Leu Arg Val Ala Val Arg Glu Gln His		
	165 170 175		
	Leu Ala Lys Pro Leu Ala Gly Val Leu Ala Val Cys Pro Val Leu Asp		
	180 185 190		
40	Pro Ala His Thr Met Met Ala Leu Asn Arg Gly Ala Phe Phe Tyr Gly		
	195 200 205		
	Arg Tyr Phe Ala His Lys Trp Lys Arg Ser Leu Thr Ala Lys Leu Ala		
	210 215 220 225		
45	Ala Phe Pro Asp Tyr Lys Tyr Gly Lys Asp Leu Lys Ser Ile His Thr		
	230 235 240		
	Leu Asp Glu Leu Asn Asn Tyr Phe Ile Pro Arg Tyr Thr Gly Phe Asn		
	245 250 255		

50

55

Ser Val Ser Glu Tyr Phe Lys Ser Tyr Thr Leu Thr Gly Gln Lys Leu
260 265 270

5 Ala Phe Leu Asn Cys Pro Ser Tyr Ile Leu Ala Ala Gly Asp Asp Pro
275 280 285

Ile Ile Pro Ala Ser Asp Phe Gln Lys Ile Ala Lys Pro Ala Asn Leu
290 295 300 305

10 His Ile Thr Val Thr Gln Gln Gly Ser His Cys Ala Tyr Leu Glu Asn
310 315 320

Leu His Lys Pro Ser Ala Ala Asp Lys Tyr Ala Val Lys Leu Phe Gly
325 330 335

15 Ala Cys

(2) INFORMATION FOR SEQ ID NO:41:

20 (i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 311 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

25 (ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

30

35

40

45

50

55

Met Leu Asp Met Pro Ile Asp Pro Val Tyr Tyr Gln Leu Ala Glu Tyr
 1 5 10 15

Phe Asp Ser Leu Pro Lys Phe Asp Gln Phe Ser Ser Ala Arg Glu Tyr
 5 20 25 30

Arg Glu Ala Ile Asn Arg Ile Tyr Glu Glu Arg Asn Arg Gln Leu Ser
 10 35 40 45

Gln His Glu Arg Val Glu Arg Val Glu Asp Arg Thr Ile Lys Gly Arg
 15 50 55 60

Asn Gly Asp Ile Arg Val Arg Val Tyr Gln Gln Lys Pro Asp Ser Pro
 20 65 70 75 80

Val Leu Val Tyr Tyr His Gly Gly Phe Val Ile Cys Ser Ile Glu
 25 85 90 95

Ser His Asp Ala Leu Cys Arg Arg Ile Ala Arg Leu Ser Asn Ser Thr
 30 100 105 110

Val Val Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Ala
 35 115 120 125

Ala Val Tyr Asp Cys Tyr Asp Ala Thr Lys Trp Val Ala Glu Asn Ala
 40 130 135 140

Glu Glu Leu Arg Ile Asp Pro Ser Lys Ile Phe Val Gly Gly Asp Ser
 45 145 150 155 160

Ala Gly Gly Asn Leu Ala Ala Ala Val Ser Ile Met Ala Arg Asp Ser
 50 165 170 175

Gly Glu Asp Phe Ile Lys His Gln Ile Leu Ile Tyr Pro Val Val Asn
 55 180 185 190

Phe Val Ala Pro Thr Pro Ser Leu Leu Glu Phe Gly Glu Gly Leu Trp
 60

35	195	200	205
	Ile Leu Asp Gln Lys Ile Met Ser Trp Phe Ser Glu Gln Tyr Phe Ser		
	210	215	230
40	Arg Glu Glu Asp Lys Phe Asn Pro Leu Ala Ser Val Ile Phe Ala Asp		
	235	240	245 250
	Leu Glu Asn Leu Pro Pro Ala Leu Ile Ile Thr Ala Glu Tyr Asp Pro		
	255	260	265
45	Leu Arg Asp Glu Gly Glu Val Phe Gly Gln Met Leu Arg Arg Ala Gly		
	270	275	280
	Val Glu Ala Ser Ile Val Arg Tyr Arg Gly Val Leu His Gly Phe Ile		
	285	290	295
50	Asn Tyr Tyr Pro Val Leu Lys Ala Ala Arg Asp Ala Ile Asn Gln Ile		
	300	305	310
	Ala Ala Leu Leu Val Phe Asp		
	315	320	

55 (2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 305 AMINO ACIDS
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

5 (ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEO ID NO:42:

10	Met Pro Leu Asp Pro Arg Ile Lys Lys Leu Leu Glu Ser Ala Leu Thr			
	5	10	15	
	Ile Pro Ile Gly Lys Ala Pro Val Glu Glu Val Arg Lys Ile Phe Arg			
	20	25	30	
15	Gln Leu Ala Ser Ala Ala Pro Lys Val Glu Val Gly Lys Val Glu Asp			
	35	40	45	
	Ile Lys Ile Pro Gly Ser Glu Thr Val Ile Asn Ala Arg Val Tyr Phe			
	50	55	60	
20	Pro Lys Ser Ser Gly Pro Tyr Gly Val Leu Val Tyr Leu His Gly Gly			
	65	70	75	80
	Gly Phe Val Ile Gly Asp Val Glu Ser Tyr Asp Pro Leu Cys Arg Ala			
	85	90	95	
25	Ile Thr Asn Ala Cys Asn Cys Val Val Val Ser Val Asp Tyr Arg Leu			
	100	105	110	
	Ala Pro Glu Tyr Lys Phe Pro Ser Ala Val Ile Asp Ser Phe Asp Ala			
	115	120	125	
30	Thr Asn Trp Val Tyr Asn Asn Leu Asp Lys Phe Asp Gly Lys Met Gly			
	130	135	140	
	Val Ala Ile Ala Gly Asp Ser Ala Gly Gly Asn Leu Ala Ala Val Val			
	145	150	155	160

35

40

45

50

55

Ala Leu Leu Ser Lys Gly Lys Ile Asn Leu Lys Tyr Gln Ile Leu Val
165 170 175

5 Tyr Pro Ala Val Ser Leu Asp Asn Val Ser Arg Ser Met Ile Glu Tyr
180 185 190

Ser Asp Gly Phe Phe Leu Thr Arg Glu His Ile Glu Trp Phe Gly Ser
195 200 205

10 Gln Tyr Leu Arg Ser Pro Ala Asp Leu Leu Asp Phe Arg Phe Ser Pro
210 215 220

Ile Leu Ala Gln Asp Phe Asn Gly Leu Pro Pro Ala Leu Ile Ile Thr
225 230 235 240

15 Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Ala Tyr Ala Asn Lys
245 250 255

Leu Leu Gln Ala Gly Val Ser Val Thr Ser Val Arg Phe Asn Asn Val
260 265 270

20 Ile His Gly Phe Leu Ser Phe Phe Pro Leu Met Glu Gln Gly Arg Asp
275 280 285

Ala Ile Gly Leu Ile Gly Ser Val Leu Arg Arg Val Phe Tyr Asp Lys
290 295 300

25 Ile
305

(2) INFORMATION FOR SEQ ID NO:43:

30 (i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 605 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

35 (ii) MOLECULE TYPE: GENOMIC DNA

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

45

50

55

EP 0 880 590 B1

	ATG AAG GTT AAA CAC GTT ATT GTT TTA CAT GGC TTA TAT ATG TCT GGC Met Lys Val Lys His Val Ile Val Leu His Gly Leu Tyr Met Ser Gly 1 5 10 15	48
5	TTC GTG ATG CGC CGG TTA TGT TCG CGT CTA GAA GAG TCG GGG GTT AAA Leu Val Met Arg Pro Leu Cys Ser Arg Leu Glu Glu Ser Gly Val Lys 20 25 30	96
	GTT TTA AAC TTA ACC TAC AAT ACT CGA GAC CCT AAT CGA GAT GCT ATT Val Leu Asn Leu Thr Tyr Asn Thr Arg Asp Pro Asn Arg Asp Ala Ile -- 35 40 45	144
10	TTT ACG CAA ATA GAT GAG TTT ATT AGC AAT GAG CCT TCT GCT TTA GTG Phe Thr Gln Ile Asp Glu Phe Ile Ser Asn Glu Pro Ser Ala Leu Val 50 55 60	192
	TGT CAC TCT ATG GGG GGC TTA GTT CCT CGC GCC TAT TTA GAG GCA AAC Cys His Ser Met Gly Gly Leu Val Ala Arg Ala Tyr Leu Glu Ala Asn 65 70 75 80	240
15	TCA GCG CCA AGT CAT CAT GTT GAA AAG GTA ATC ACC TTA GGA ACG CCA Ser Ala Pro Ser His His Val Glu Lys Val Ile Thr Leu Gly Thr Pro 85 90 95	288
	CAT ACT GGC AUC CAT ATT GCT GAA AAA ATG CAG CAA AAA GGG TTC GAG His Thr Gly Ser His Ile Ala Glu Lys Met Gln Gln Lys Gly Phe Glu 100 105 110	336
20		
25	CTA TTA TTA AAA AAT AGC GTT GAG TTT TTA CTC TCT AAG AAT GGT GAT Leu Leu Leu Lys Asn Ser Val Glu Phe Leu Leu Ser Lys Asn Gly Asp 115 120 125	384
	TGG CCT TTT AAA GCC AAG CTA TAT AGC ATT GCC ATT GGC GAC TTA CCG ATT Trp Pro Phe Lys Ala Lys Leu Tyr Ser Ile Ala Gly Asp Leu Pro Ile 130 135 140	432
30	GGC TTA ATG CCA CTC ATT GTA AAA CGC ACC CGC TCT GAT GGC ACT GTA Gly Leu Met Pro Leu Ile Val Lys Gly Ser Arg Ser Asp Gly Thr Val 145 150 155 160	480
	TTC CTA GAT GAA ACC AAG CTA AAG GGT ATG GCT GAA CAC ARG GTG TTT Leu Leu Asp Glu Thr Lys Leu Lys Gly Met Ala Glu His Lys Val Phe 165 170 175	528
35	CAT TTA AGC CAT ACA AGT ATG ATT TAC TCT CGC CAA GTC GTT AAT TAT His Leu Ser His Thr Ser Met Ile Tyr Ser Arg Gln Val Val Asn Tyr 180 185 190	576
	ATT CTT GAG CGC TTG AAC GAG GAC ATT TA Ile Leu Glu Arg Leu Asn Glu Asp Ile 195 200	605
40		

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS

- 45 (A) LENGTH: 779 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

50 (ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

EP 0 880 590 B1

	ATG ATA AAA AAC TTC GAC AGA GAA AAT TCT AGC TTA GTC CTG TCC GGT Met Ile Lys Asn Phe Asp Arg Glu Asn Ser Ser Leu Val Leu Ser Gly 1 5 10 15	48
5	GGT GGT CCT CTG GGT ATT GCT CAC TTG GGT GTA CTG CAT GAC CTT GAA Gly Gly Ala Leu Gly Ile Ala His Leu Gly Val Leu His Asp Leu Glu 20 25 30	96
	AAA CAA AAT ATT GTC CCA AAT GAA ATT GTC GGT ACA AGT ATG GGT GGT Lys Gln Asn Ile Val Pro Asn Ile Val Gly Thr Ser Met Gly Gly 35 40 45	144
10	ATC ATT GGT GCA TCT ATG GCT ATC GGG ATG AAA GAG AAA GAA ATA CTC Ile Ile Gly Ala Ser Met Ala Ile Gly Met Lys Glu Lys Glu Ile Leu 50 55 60	192
	GAA GAA ATC AAA AAC TTT TCC AAT GTC TTC AAC TGG ATA AAA TTC TCT Glu Glu Ile Lys Asn Phe Ser Asn Val Phe Asn Trp Ile Lys Phe Ser 65 70 75 80	240
15	TTT TCC GGT AAT TCT GTT GTC GAT AAC GAG AAG ATC GCT AAG ATA TTT Phe Ser Gly Asn Ser Val Val Asp Asn Glu Lys Ile Ala Lys Ile Phe 85 90 95	288
	GAT ACT CTT TTT AAA GAC AGA AAG ATG ACA GAT ACG GTG ATC CCT CTT Asp Thr Leu Phe Lys Asp Arg Lys Met Thr Asp Thr Val Ile Pro Leu 100 105 110	336
20	AAA CTC ATC GCT ACA AAC TTA CAT AAT GGA CAT AAA AAA GTC TTT ACT Lys Leu Ile Ala Thr Asn Leu His Asn Gly His Lys Lys Val Phe Thr 115 120 125	384
	GCT TCG GAT GAT GTC CTG ATC AAA GAT GCA ATA CTC TCA ACA ATG GCA Ala Ser Asp Asp Val Leu Ile Lys Asp Ala Ile Leu Ser Thr Met Ala 130 135 140	432
25	ATA CCC GGT GTC TTT GAA GAA CAT ATT ATT GAT GGT GAA ACC TAT GGC Ile Pro Gly Val Phe Glu Asn Ile Ile Asp Gly Glu Thr Tyr Gly 145 150 155 160	480
30	GAC GGT TTT CTT TGT GAA AAC CTT GGT GTG AAT GAG GCA ACA TTC AAT Asp Gly Phe Leu Cys Glu Asn Leu Gly Val Asn Glu Ala Thr Phe Asn 165 170 175	528
35	GAT GTT TTA GCT GTC GAT GTC ATG GGT GAG AAC TCT TTT GAA AAA GCA Asp Val Leu Ala Val Asp Val Met Gly Glu Asn Ser Phe Glu Lys Ala 180 185 190	576
	ATG CCG GAC AAC TTC TTT AAA ACA TCA AAT GTT TTA GAA ATG TTT GAA Met Pro Asp Asn Phe Phe Lys Thr Ser Asn Val Leu Glu Met Phe Glu 195 200 205	624
40	AAA TCA ATG CGA CTT TTT ATT TAC AAC CAG ACA CAG ACA CAT ATT AAA Lys Ser Met Arg Leu Phe Ile Tyr Asn Gln Thr Gln Thr His Ile Lys 210 215 220	672
	AAT GCA AAT AAA AAT ATT TAT CTT ATT GAA CCC GTT ACC AAA GAG TAT Asn Ala Asn Lys Ile Tyr Leu Ile Glu Pro Val Thr Lys Glu Tyr 225 230 235 240	720
45	AAA ACA TTT CRA TTT CAT AAA CAT AAA GAG ATA CGT GCT TTA GGC TTG Lys Thr Phe Gln Phe His Lys Lys Glu Ile Arg Ala Leu Gly Leu 245 250 255	768
	GCT TTA CTG TG Gly Leu Leu	779
50	(2) INFORMATION FOR SEQ ID NO:45:	

(i) SEQUENCE CHARACTERISTICS

- 55 (A) LENGTH: 905 NUCLEOTIDES
 (B) TYPE: NUCLEIC ACID
 (C) STRANDEDNESS: SINGLE
 (D) TOPOLOGY: LINEAR

EP 0 880 590 B1

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

5	ATG CCC CTA CAT CCA AAG GTC AAG AAA TTA CTT TCC CAG CTA CCT CCC Met Pro Leu His Pro Lys Val Lys Lys Leu Leu Ser Gln Leu Pro Pro 1 5 10 15	48
10	CAG GAC TTC TCC AGA AAC GTG CAG GAC CTG AGC AAG GCC TGG GAT TTA Gln Asp Phe Ser Arg Asn Val Gln Asp Leu Arg Lys Ala Trp Asp Leu 20 25 30	96
	CCC TTC TCA GGG AGG AGG GAG ACC CTG AAG AGG GTT GAG GAC CTT GAG Pro Phe Ser Gly Arg Arg Glu Thr Leu Lys Arg Val Glu Asp Leu Glu 35 40 45	144
15	ATA CCC ACT AGG GAC GCA CGA ATC AGG GCC AGG GTC TAC ACC CCC TCA Ile Pro Thr Arg Asp Ala Arg Ile Arg Ala Arg Val Tyr Thr Pro Ser 50 55 60	192
	AGT AAG GAA AAC TTA CCC GTC CTT GTT TAC TAT CAC GGC GGT GGC TTC Ser Lys Glu Asn Leu Pro Val Leu Val Tyr Tyr His Gly Gly Gly Phe 65 70 75 80	240
20	GTC TTC GGT AGC GTT GAC AGC TAC GAC GGC CTC GCA TCC CTT ATT GCC Val Phe Gly Ser Val Asp Ser Tyr Asp Gly Leu Ala Ser Leu Ile Ala 85 90 95	288
	AAG GAA TCT GGG ATT GCG GTT ATC TCC GTG GAG TAT AGG CTC GCC CCT Lys Glu Ser Gly Ile Ala Val Ile Ser Val Glu Tyr Arg Leu Ala Pro 100 105 110	336
25	GAG CAC AAG TTC CCC ACC GCA GTC AAC GAC TCG TGG GAT GCG CTT CTC Glu His Lys Phe Pro Thr Ala Val Asn Asp Ser Trp Asp Ala Leu Leu 115 120 125	384
	TGG ATC GCG GAG AAC GGA GGC AAG CTG GGG CTC GAC ACC TCG AGA CTT Trp Ile Ala Glu Asn Gly Gly Lys Leu Gly Leu Asp Thr Ser Arg Leu 130 135 140	432
30	GCC GTG GCT GGG GAT AGT GCT GGA GGA AAC CTG TCT GCC GTG GTG TCC Ala Val Ala Gly Asp Ser Ala Gly Gly Asn Leu Ser Ala Val Val Ser 145 150 155 160	480

35

40

45

50

55

EP 0 880 590 B1

	CTC CTG GAC AGG GAC CAG GGT AAG GGA CTG GTT AGT TAT CAG GTC CTA Leu Leu Asp Arg Asp Gln Gly Lys Gly Leu Val Ser Tyr Gln Val Leu 165 170 175	528
5	TAC TAC CCA GCA GTG AAC ATG GTC GAT AAC TCC CCA TCC GTC AGG GAG Ile Tyr Pro Ala Val Asn Met Val Asp Asn Ser Pro Ser Val Arg Glu 180 185 190	576
	TAC GGC GAG GGA TAC TTC CTC ACC AGG TCC ATG ATG AAC TGG TTC GGG Tyr Gly Glu Gly Tyr Phe Leu Thr Arg Ser Met Met Asn Trp Phe Gly 195 200 205	624
10	ACC ATG TAC TTC TCC TCT GGA AGG GAA GCG GTC TCC CCC TAC GCC TCT Thr Met Tyr Phe Ser Ser Gly Arg Glu Ala Val Ser Pro Tyr Ala Ser 210 215 220	672
	CCA GCC TTG GCT GAC CTA CAT AAC CTC CCA CCC TCA CTG GTG ATC ACT Pro Ala Leu Ala Asp Leu His Asn Leu Pro Pro Ser Leu Val Ile Thr 225 230 235 240	720
15	GCA GAG TAT GAT CCC CTA AGG GAT CAG GGA GAG ACC TAC TCT CAC TCC Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Thr Tyr Ser His Ser 245 250 255	768
20	CTA AAC GAG GCT GGA AAC GTC TCA ACC TTG GTT AGA TAT CAA GGA ATG Leu Asn Glu Ala Gly Asn Val Ser Thr Leu Val Arg Tyr Gln Gly Met 260 265 270	816
	ATT CAC GGC TTC CTG TCC TAC GAG TGG ATA ACT GCC GGT AAA CTA Ile His Gly Phe Leu Ser Phe Tyr Glu Trp Ile Thr Ala Gly Lys Leu 275 280 285	864
25	GCC ATT CAC CAC ATT GCT GGG GTT CTG AGA TCT GTC CTT TA Ala Ile His His Ile Ala Gly Val Lys Arg Ser Val Leu Arg Ser Val 290 295 300	905
	Leu	301

(2) INFORMATION FOR SEQ ID NO:46:

- 30 (i) SEQUENCE CHARACTERISTICS
- 35 (A) LENGTH: 978 NUCLEOTIDES
 - (B) TYPE: NUCLEIC ACID
 - (C) STRANDEDNESS: SINGLE
 - (D) TOPOLOGY: LINEAR
- 40 (ii) MOLECULE TYPE: GENOMIC DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

45

50

EP 0 880 590 B1

	GTC GCC TTC TTC GAT ATG CCC CTT GAG GAA CTG AAA AAG TAC CGG CCT Val Ala Phe Phe Asp Met Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro 3 5 10 15	48
5	GAA AGG TAC GAG GAG AAA GAT TTC GAT GAG TTC TGG AGG GAA ACA CTT Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Arg Glu Thr Leu 20 25 30	96
	AAA GAA AGC GAA GGA TTC CCT CTG GAT CCC GTC TTT GAA AAG GTG GAC Lys Glu Ser Glu Gly Phe Pro Leu Asp Pro Val Phe Glu Lys Val Asp 35 40 45	144
10	TTC CAT CTC AAA ACG GTT GAA ACG TAC GAT GTT ACT TTC TCT GGA TAC Phe His Leu Lys Thr Val Glu Thr Tyr Asp Val Thr Phe Ser Gly Tyr 50 55 60	192
	ACG GGG CAG AGA ATA AAG GGC TGG CTT CTT GTT CCG RAG TTG GCG GAA Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Ala Glu 65 70 75 80	240
15	GAA AAG CTT CCA TCC GTC GTG CAG TAC ATA GGT TAC AAT GOT GGA AGG Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg 85 90 95	288
	GGT TTT CCR CAC GAC TGG CTG TTC TGG CCG TCA ATG GOT TAC ATC TGT Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys 100 105 110	336

20

25

30

35

40

45

50

55

EP 0 880 590 B1

	TTT GTC ATG GAC ACC AGG GGG CAG GGA AGC GGC TGG ATG AAG GGA GAC Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Met Lys Gly Asp 115 120 125	384
5	ACA CCG GAT TAC CCT GAG GGT CCA GTC GAT CCA CAG TAC CCC GGA TTC Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe 130 135 140	432
	ATG ACG AGG GGC ATT CTG GAT CCG GGA ACC TAT TAC TAC AGG CGA GTC Met Thr Arg Gly Ile Leu Asp Pro Gly Thr Tyr Tyr Arg Arg Val 145 150 155 160	480
10	TTC GTG GAT GCG GTC AGG GCG GTG GAA GCA GCC ATT TCC TTC CCG AGA Phe Val Asp Ala Val Arg Ala Val Glu Ala Ala Ile Ser Phe Pro Arg 165 170 175	528
	GTC GAT TCC AGG AAG GTC GTG GTG CCC GGA GGC AGT CAG GGT GGG GGA Val Asp Ser Arg Lys Val Val Val Ala Gly Gly Ser Gln Gly Gly Gly 180 185 190	576
15	ATC CCC CTT GCG GTG AGT GCC CTG TCG AAC AGG GTC AAG GCT CTG CTC Ile Pro Leu Ala Val Ser Ala Leu Ser Asn Arg Val Lys Ala Leu Leu 195 200 205	624
	TGC GAT GTG CCG TTT CTG TGC CAC TTC AGA AGG GCC GTG CAA CTT GTC Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val 210 215 220	672
20	GAC ACA CAC CCA TAC GTG GAG ATC ACC AAC TTC CTC AAA ACC CAC AGG Asp Thr His Pro Tyr Val Glu Ile Thr Asn Phe Leu Lys Thr His Arg 225 230 235 240	720
	GAC AAA GAG GAG ATT GTT TTC AGA ACA CTT TCC TAC TTC GAT GGT GTG Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val 245 250 255	768
25	AAC TTT GCA GCA AGG GCA ARG GTG CCC GGC CTG TTT TCC GTT GGG CTC Asn Phe Ala Ala Arg Ala Lys Val Pro Ala Leu Phe Ser Val Gly Leu 260 265 270	816
	ATG GAC ACC ATC TGT CCT CCC TCG ACG GTC TTC GCC GCT TAC AAC CAC Met Asp Thr Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn His 275 280 285	864
30	TAC GCC GGT CCA AAG GAG ATC AGA ATC TAT CCG TAC AAC AAC CAC GAA Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu 290 295 300	912
	GGT GGA GGT TCT TTC CAG GCA ATT GAG CAG GTG AAA TTC TTG AAG AGA Gly Gly Gly Ser Phe Gln Ala Ile Glu Gln Val Lys Phe Leu Lys Arg 305 310 315 320	960
35	CTA TTT GAG GAA GGC TAG Leu Phe Glu Glu Gly 325	978

40

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS

- 45 (A) LENGTH: 879 NUCLEOTIDES
 (B) TYPE: NUCLEIC ACID
 (C) STRANDEDNESS: SINGLE
 (D) TOPOLOGY: LINEAR

50 (ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

55

EP 0 880 590 B1

	ATG CGC ACC CTC TCC TTC GGT CGG ATG ACC ACA CGG GGA AGC ATT CAC Met Arg Thr Leu Ser Phe Gly Pro Met Thr Thr Gly Gly Ser Ile His 1 5 10 15	48
5	ATG GCG ACC ATG GAC GTG ATG CGC GGG CCC GGG ATG CAG CGG CTG TCA Met Ala Thr Met Asp Val Met Arg Gly Pro Gly Met Gin Arg Leu Ser 20 25 30	96
10	CAG GGC GCC AGG GAG GGC AAC CAC CCC TGG GCG AAG CGA CTG GGC Gln Gly Ala Arg Glu Ala Ala Asn His Pro Trp Ala Lys Arg Leu Gly 35 40 45	144
15	CGC ATG GGC TAC GCG GCC AAG GGC CCT GTG TAC GCC ATC ATC GGC GTG Arg Met Gly Tyr Ala Ala Lys Gly Ala Val Tyr Ala Ile Ile Gly Val 50 55 60	192
20	CTC GCG CTG AAG CTC CGG CGG CGG CGG CGG CGG ACC ACG GAC AGC Leu Ala Leu Lys Leu Ala Ala Gly Glu Gly Gly Arg Thr Thr Asp Ser 65 70 75 80	240
25	CAC GGC GCG GTG AAC ACC GTG CGG CAC CGG CCC TTC GGC GTC GCG CTG His Gly Ala Val Asn Thr Val Ala His Gly Pro Phe Gly Val Ala Leu 85 90 95	288
30	CTG GCG CTG CTG CTG CTG CGC CTG CTG GGC TAC GTG GTC TGG AGG TTC Leu Ala Val Leu Val Val Gly Leu Leu Gly Tyr Val Val Trp Arg Phe 100 105 110	336
35	GCC CAG GGC TTC GTG GAC ACG GAG GAC AAG GGC TCC GAC GCG AAG CGA Ala Gln Ala Phe Val Asp Thr Glu Asp Lys Gly Ser Asp Ala Lys Gly 115 120 125	384
40	ATC GCC ACG CGC GCT ATG TAC TTC CTC AGC GGC TGC ATC TAC CGG TCG Ile Ala Thr Arg Ala Met Tyr Phe Leu Ser Gly Cys Ile Tyr Ala Ser 130 135 140	432
45	CTG GCC TTC TTC GGC CGG CAG TCC CTG GTC GGC GCG CAC GGC CGG Leu Ala Phe Phe Ala Ala Gln Ser Leu Val Gly Ala Ala His Gly Arg 145 150 155 160	480
50	AGC AAG GGG ACG CAG CGC TGG ACG GCC ACC CTG ATG GAG CAG CCC TTT Ser Lys Gly Thr Gln Gly Trp Thr Ala Thr Leu Met Glu Gln Pro Phe 165 170 175	528
55	GGC CGC GTG CTG CTG CGG CTG CTG GGG CTG GGC ATC GTG GCG TTC GCG Gly Arg Val Leu Val Ala Leu Val Gly Leu Gly Ile Val Gly Phe Ala 180 185 190	576
60	CTG AAG CAG TTC CAC ACC GCG TGG AAG CGG AAG TTC CCG GAG AAG CTC Leu Lys Gln Phe His Thr Ala Trp Lys Ala Lys Phe Arg Glu Lys Leu 195 200 205	624
65	ACC CTC ACC GGA CTG GCT GCC CGG AAG CAG CAC CAC ATC GAG CGC ATG Thr Leu Thr Gly Leu Ala Ala Arg Lys Gln His His Ile Glu Arg Met 210 215 220	672
70	TGC CAG TTC GGC ATC GCC CGG CGC CGC GTC GTG TTC GGC GTC ATC GGC Cys Gln Phe Gly Ile Ala Ala Arg Gly Val Val Phe Ala Val Ile Gly 225 230 235 240	720
75	GCG TTC CTC GTC CGC TCC GGC GTC GAC CGG AAC CCC CGG GAG GGC AAG Gly Phe Leu Val Arg Ser Ala Val Asp Ala Asn Pro Gly Glu Ala Lys 245 250 255	768
80	GGC CTG GGA GAG GCC CTG GCC GTC GTC CGG AGG CAG CGG TCC GGC GAC Gly Leu Gly Glu Ala Leu Ala Val Val Ala Arg Gln Pro Ser Gly Asp 260 265 270	816
85	GTG CTC CTG CGG GTG GTC GCG CGG CGC CTG GTG CCC TAC GCC GCG TAC Val Leu Leu Gly Val Val Ala Ala Gly Leu Val Ala Tyr Ala Ala Tyr 275 280 285	864
90	CTG TTC CTC CAG CGG CGC TAC CGC GAA CTC TAG Leu Phe Leu Gln Ala Arg Tyr Arg Glu Leu 290 295	897

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 914 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- 5 (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

10

15

20

25

30

35

40

45

50

55

EP 0 880 590 B1

(xi) SEQUENCE DESCRIPTION: SEQ-ID NO:48:

	ATG AGC AAA TTC GCA ATA CTC TGG GCG TTG ATA ACG GCA TAC CTG CCG Met Ser Lys Phe Ala Ile Leu Trp Ala Leu Ile Thr Ala Tyr Leu Pro 1 5 10 15	48
5	GAA CCT GTG ATG AAA CTG GTA TAT TTA GCG CGG CGC GAA ACG CTT GGG Glu Pro Val Met Lys Leu Val Tyr Leu Gly Arg Arg Glu Thr Leu Gly 20 25 30	96
10	GCA CGG ACG CTT GAC GTT AAA GCC CAA CCT GTC GGG CGG CTG GCC AAT Ala Arg Thr Leu Asp Val Lys Ala Gln Ala Val Gly Arg Leu Ala Asn 35 40 45	144
15	GCA ACA AGA CCT GTC GGG GTG ATT CCG ACG GTC GAG GAA AGC CGG AAG Ala Thr Arg Pro Val Gly Val Ile Pro Thr Val Glu Glu Ser Arg Lys 50 55 60	192
20	ATG ACG GAT AAA GCC GTT AGC CTT TTT GAT CAG CCC GCC CCC GAA TTA Met Thr Asp Lys Ala Val Ser Leu Phe Asp Gln Pro Ala Pro Glu Leu 65 70 75 80	240
25	TTC CGT AAA AAA GAC ATT CAG ATT GAC GCG CCT GAA GGG CCT ATT GAT Phe Arg Lys Lys Asp Ile Gln Ile Asp Gly Ala Glu Gly Pro Ile Asp 95 90 95	288
30	GCC CGT ATT TAC AGC GGC CCT GCA AAA CAT CGC CCA CCA CCA ATA CTA Ala Arg Ile Tyr Ser Gly Pro Ala Lys His Arg Pro Arg Pro Ile Leu 100 105 110	336
35	GTC TAT TTT CAC GGC GGT GGC TGG GTT CAG GGC AAT CTG GAC AGC CAT Val Tyr Phe His Gly Gly Trp Val Gln Gly Asn Leu Asp Ser His 115 120 125	384
40	GAC GGG GTT TGC GGC AAG CTG GCA AAA TGG GCG AAC TGC ATT GTT ATC Asp Gly Val Cys Gly Lys Leu Ala Lys Trp Ala Asn Cys Ile Val Ile 130 135 140	432
45	TCG GTC GAT TAT CGT CTA GCG CCC GAA CAC AAA TTT CCT TGT GCG CGG Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Cys Ala Pro 145 150 155 160	480
50	CTT GAT GCG ATT GCG GCC TAT AAA TGG GTG CGC GCC AAC GCA ACA AAC Leu Asp Ala Ile Ala Ala Tyr Lys Trp Val Arg Ala Asn Ala Thr Asn 165 170 175	528
55	CTT GGC GGC GAT CCT GAA CGT ATC GGC GTT GGC GGC GAT AGC GCA GCG Leu Gly Gly Asp Pro Glu Arg Ile Gly Val Gly Asp Ser Ala Gly 180 185 190	576
60	GGC AAT CTT CCC GCC GTT GTC TGC CAA CAA ACC GCC ATG AAC GGC GAG Gly Asn Leu Ala Ala Val Val Cys Gln Gln Thr Ala Met Asn Gly Glu 195 200 205	624
65	CGC ACA CCA GAT CTG CAA GTC CTG ATC TAT CCG GCG CTG GAT GCA CGC Arg Thr Pro Asp Leu Gln Val Leu Ile Tyr Pro Ala Leu Asp Ala Arg 210 215 220	672
70	ATG ATC TCG ACC TCG ATG GAG GAA TTG CGT GAT GCC TAC ATC TTG CCG Met Ile Ser Thr Ser Met Glu Glu Leu Arg Asp Ala Tyr Ile Leu Pro 225 230 235 240	720
75	AAA TCC AGA ATG GAG TAT TTC CTC GGC CTA TAT ACU CGT GGC CCT GAC Lys Ser Arg Met Glu Tyr Phe Leu Gly Leu Tyr Thr Arg Gly Pro Asp 245 250 255	768
80	GAT ATC GAG GAC CTT AGG ATG TCG CCA ATT CTC ACG GAT ACC GTC GCG Asp Ile Glu Asp Leu Arg Met Ser Pro Ile Leu Arg Asp Thr Val Ala 260 265 270	816
85	GAT CAA CCC CAA GCC TGC ATT GTC ACC TGT GGG TTG GAC CCT GCG CCA Asp Gln Pro Gln Ala Cys Ile Val Thr Cys Gly Phe Asp Pro Ala Arg 275 280 285	864
90	CGA CGG GAA CAC CTA CGC CGA ACG CTT ATT TGC CGA GGG GAT AGA CGT Arg Arg Glu His Leu Arg Arg Thr Leu Asn Cys Arg Gly Asp Arg Arg 290 295 300	912
95	TA	914

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 926 NDCLEOTIDES
 (B) TYPE: NUCLSIC ACID
 5 (C) STRANDEDNESS: SINGLE
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

	GTC AGC ATT CGT CTG CGA CTG TTA AAC TGG TTT TTG AAT ACC TTT GAA	48
5	Val Ser Ile Arg Leu Arg Leu Leu Asn Trp Phe Leu Asn Thr Phe Glu	
	1 5 10 15	
15	AAA CCA AAA CTG GCC GCG AAA ACG CCG GAT TTG CGA AAA TCG	96
	Lys Pro Lys Leu Ala Ala Lys Thr Pro Asp Asp Leu Arg Lys Ser	
	20 25 30	
20	TTT GAA TTA AAG GCG AGG TTT TTG TTT CCG GCG CCA CGT AAA ACA AGG	144
	Phe Glu Leu Lys Ala Arg Phe Leu Phe Pro Ala Pro Arg Lys Thr Arg	
	35 40 45	
25	TTT AGT CAT GAT GTA TTG CAG TCA GGC ATC GGG TCG GTA AAT GCC CGG	192
	Phe Ser His Asp Val Leu Gln Ser Gly Ile Gly Ser Val Asn Asn Gln	
	50 55 60	
30	TGG GCG AAA TCC AAA TCT GCA TCT GAT GAC AGG GTA ATC CTG TAT TTT	240
	Trp Ala Lys Ser Ala Ser Asp Asp Arg Val Ile Leu Tyr Phe	
	65 70 75 80	
	CAT GGG GGA GGG TAT GTT TTT GGG TCA CCA AAA ACG CAC CGT GCA ATG	288
	His Gly Gly Tyr Val Phe Gly Ser Pro Lys Thr His Arg Ala Met	
	85 90 95	
35	TTG GCG CGC TTG TCG GCA ATG AAT GOT CTT TCT GCG TGC CTT CCA GAT	336
	Leu Ala Arg Leu Ser Ala Met Thr Gly Leu Ser Ala Cys Leu Pro Asp	
	100 105 110	
	TAT AGG TTG GCA CCA GAG CAC CCA TTT CCA GCC GCG ATC GAA GAT GCA	384
	Tyr Arg Leu Ala Pro Glu His Pro Phe Pro Ala Ala Ile Glu Asp Ala	
	115 120 125	
40	GTT TTA TCG TAT AAA TGT TTA CTA GAG CGA GCA ATC GAG CCC CAA AAT	432
	Val Leu Ser Tyr Lys Cys Leu Leu Glu Arg Ala Ile Glu Pro Gln Asn	
	130 135 140	
	ATT ATA CTG GGG GGG GAC AGT GCT GGT GGC GGT TTG GTT CTT GCT TTG	480
	Ile Ile Leu Gly Gly Asp Ser Ala Gly Gly Leu Val Leu Ala Leu	
	145 150 155 160	
45	CTT GCA GAA ATC AMG GCC CAA TCC TTG CCT AAA CCT GCT GGC GTT TTT	528
	Leu Ala Glu Ile Lys Ala Gln Ser Leu Pro Lys Pro Ala Gly Val Phe	
	165 170 175	
	GCC TTG TCG CCT TTG GTT GAT TTA TCA TTT TCG GGC CTT TCG TTT TCT	576
	Ala Leu Ser Pro Leu Val Asp Leu Ser Phe Ser Gly Leu Ser Phe Ser	
	180 185 190	
50	AAA AAT GCC CAA ACC GAT GTG ATG TTG CCC GCA TCA CGG GCT GCG GAT	624
	Lys Asn Ala Gln Thr Asp Val Met Leu Pro Ala Ser Arg Ala Ala Asp	
	195 200 205	
	ATG GCG ACC TTG TAT TTG GAT GGG GCG GAT GCA GAT GAT CCA CGT GCA	672
	Met Ala Thr Leu Tyr Leu Asp Gly Ala Asp Ala Asp Asp Pro Arg Ala	
	210 215 220	
55	TCG CCG CTG CAG GCG GAT TTT TCT GGC ATG CCG CCT GCA TTT CTG ACA	720
	Ser Pro Leu Gln Ala Asp Phe Ser Gly Met Pro Pro Val Phe Leu Thr	
	225 230 235 240	
	GCA AGT GAC AGT GAA ATC CTG TTG GAT GAT TGC CTG CGG ATG GCG GAT	768
	Ala Ser Asp Ser Glu Ile Leu Leu Asp Asp Cys Leu Arg Met Ala Asp	
	245 250 255	
	CAC TTG CGT GCG CAA GGT GTC GTG ACA GAC CGG ATT GTT GAA AAC	816
	His Leu Arg Ala Gln Gly Val Val Thr Asp Arg Ile val Glu Asn	
	260 265 270	

EP 0 880 590 B1

CAT CCA CAT GTT TGG CAT ATT TTT CAA CGC CTT CTA CCC GAA GCA GAT His Pro His Val Trp His Ile Phe Gln Arg Leu Leu Pro Glu Ala Asp 275 280 285	864
5 CAG GGG CTG CGG GCG ATT GCC GCG TGG ATT AAA CCT CTT TTA TCA GGT Gln Gly Leu Arg Ala Ile Ala Ala Trp Ile Lys Pro Leu Leu Ser Gly 290 295 300	912
TCA AAC GAA AGC TA Ser Asn Glu Ser 305	926

10

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS

15

- (A) LENGTH: 713 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

20

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

25

30

35

40

45

50

55

EP 0 880 590 B1

	ATG CTT ACA TTT AAT GTT TTA TAT GGT ATG ATG AAA CAA AAA CTA GCA Met Leu Thr Phe Asn Val Leu Tyr Gly Met Met Lys Gln Lys Leu Ala 1 5 10 15	48
5	GCA ATT CTC ATG TTT TTA GGG CTA TCA GCA GCA GAG GCT CAA GAC TGG Ala Ile Leu Met Phe Leu Gly Leu Ser Ala Ala Glu Ala Gln Asp Trp 20 25 30	96
	CCT GAC CTA CAG AAA TAT CGT AGT GCT AAT AAA GAA GCC AAA TTA CTT Pro Asp Leu Gln Lys Tyr Arg Ser Ala Asn Lys Glu Ala Lys Leu Leu 35 40 45	144
10	CCA AAG GAA AAC CGG AAG GTG GTT TTT ATG GGC AAC TCC ATT ACA GAA Pro Lys Glu Asn Arg Lys Val Val Phe Met Gly Asn Ser Ile Thr Glu 50 55 60	192
	GCC TGG ATT AGT CAG CGA CCT GAG TTT TTT AGT GAA AAT GGG TTT ATC Ala Trp Ile Ser Gln Arg Pro Glu Phe Phe Ser Glu Asn Gly Phe Ile 65 70 75 80	240
15	GGT CGA GGC ATC AGT GGC CAG ACA ACC CCT CAG ATG TTG TTG AGA TTC Gly Arg Gly Ile Ser Gly Gln Thr Thr Pro Gln Met Leu Leu Arg Phe 85 90 95	288
	CGA CAG GAT GTG ATA GAC CTG CAG CCA AAG CCT GTC GTG ATA CTA GCT Arg Gln Asp Val Ile Asp Leu Gln Pro Lys Ala Val Val Ile Leu Ala 100 105 110	336
20	GGT ACC AAT GAC GTC GCT CAA AAT ACC GGG CCG ATG ACC ATT GAG GAA Gly Thr Asn Asp Val Ala Gln Asn Thr Gly Pro Met Thr Ile Glu Glu 115 120 125	384
	TCG CTT GCT AAC ATT AAG TCT ATG GTG GAG CTG GCG CAA GCC AAT GGG Ser Leu Ala Asn Ile Lys Ser Met Val Glu Leu Ala Gln Ala Asn Gly 130 135 140	432
25	ATC ACG CCT GTT TTG TGT ACC GTG CTG CCT GCA GAT CGT TTC AGC TGG Ile Thr Pro Val Leu Cys Thr Val Leu Pro Ala Asp Arg Phe Ser Trp 145 150 155 160	480
	CGA CCT GAG CTT ACA CCC GCA GAA ACT ATC ATT GCC CTC AAT CAG CTC Arg Pro Glu Leu Thr Pro Ala Glu Thr Ile Ile Ala Leu Asn Gln Leu 165 170 175	528
30	ATT AAG CAA TAT GCC GAG GCA CAG GGC CTG GCC CTG GTG GAT TAT CAT Ile Lys Gln Tyr Ala Glu Ala Gln Gly Leu Ala Leu Val Asp Tyr His 180 185 190	576
	GCT GCA CTC ACC AAT AAA GGT GGA GGA CTT CCG GTG AAA TAC GCA GAA Ala Ala Leu Thr Asn Lys Gly Gly Leu Pro Val Lys Tyr Gly Glu 195 200 205	624
35	GAT GGT GTG CAT CCA AAT GTC GCA GGC TAT CAG GTG ATG GAA AAC ATT Asp Gly Val His Pro Asn Val Ala Gly Tyr Gln Val Met Glu Asn Ile 210 215 220	672
	GTT TTA CCG GTC ATT TCC AGC GAG TTG GCA AAG CTG AAG TA Val Leu Pro Val Ile Ser Ser Glu Leu Ala Lys Leu Lys 225 230 235	713
40		
45		

(2) INFORMATION FOR SEQ ID NO:51:

50 (i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 978 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

55 (ii) MOLECULE TYPE: GENOMIC DNA

EP 0 880 590 B1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

5	ATG GCC TTC TTC GAT TTA CCA CTC GAA GAA CTG AAG AAA TAT CGT CCA Met Ala Phe Asp Leu Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro 1 5 10 15	48
	GAG CGG TAC GAA GAG AAA GAC TTC GAT GAG TTC TGG GAA GAG ACA CTC Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Glu Glu Thr Leu 20 25 30	96
10	GCA GAG AGC GAA AAG TTC CCC TTA GAC CCC GTC TTC GAG AGG ATG GAG Ala Glu Ser Glu Lys Phe Pro Leu Asp Pro Val Phe Glu Arg Met Glu 35 40 45	144
	TCT CAC CTC AAA ACA GTC GAA GCG TAC GAT GTC ACC TTC TCC GGA TAC Ser His Leu Lys Thr Val Ala Tyr Asp Val Thr Phe Ser Gly Tyr 50 55 60	192
15	AGG GGA CAG AGG ATC AAA GGG TCG CTC CTT GTT CCA AAA CTG GAA GAA Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Glu Glu 65 70 75 80	240
	GAA AAA CCT CCC TGC GTT GTG CAG TAC ATA GGA TAC AAC GGT GGA AGA Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg 85 90 95	288
20	GGA TTC CCT CAC GAC TGG CTG TTC TGG CCT TCT ATG GGT TAC ATA TGT Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys 100 105 110	336
	TTC GTC ATG GAT ACT CGA GGT CAG GGA AGC GGC TGG CTG AAA GGA GAC Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Leu Lys Gly Asp 115 120 125	384
25	ACA CGG GAT TAC CCT GAG GGT CCC GTT GAC CCT CAG TAT CCA GGA TTC Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe 130 135 140	432
	ATG ACA AGA GGA ATA CTG GAT CCC AGA ACT TAC TAC TAC AGA CGA GTC Met Thr Arg Gly Ile Leu Asp Pro Arg Thr Tyr Tyr Arg Arg Val 145 150 155 160	480
30	TTC ACG GAC GCT GTC AGA GCC GTT GAA GCT GCT GCT TCT TTT CCT CAG Phe Thr Asp Ala Val Arg Ala Val Glu Ala Ala Ala Ser Phe Pro Gln 165 170 175	538
	GTA GAT CAA GAA AGA ATC GTG ATA GCT GGA GGC AGT CAG GGT GGC GGA Val Asp Gln Glu Arg Ile Val Ile Ala Gly Gly Ser Gln Gly Gly Gly 180 185 190	576
35	ATA GCC CTT GCG GTG AGC GCT CTC TCA AAG AAA GCA AAG GCT CTT CTG Ile Ala Leu Ala Val Ser Ala Leu Ser Lys Lys Ala Lys Ala Leu Leu 195 200 205	624
	TGC GAT GTG CCG TTT CTG TGT CAC TTC AGA AGA GCA GTA CAG CTT GTG Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val 210 215 220	672

45

50

55

	GAT ACG CAT CCA TAC GCG GAG ATC ACG AAC TTT CTA AAG ACC CAC AGA Asp Thr His Pro Tyr Ala Glu Ile Thr Asn Phe Leu Lys Thr His Arg 225 230 235 240	720
5	GAC AAG GAA GAA ATC GTG TTC AGG ACT CCT TCC TAT TTC GAT GGA GTG Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val 245 250 255	768
	AAC TTC GCA GCC AGA GCG AAG ATC CCT GCG CTG TTT TCT GTG GGT CTC Asn Phe Ala Ala Arg Ala Lys Ile Pro Ala Leu Phe Ser Val Gly Leu 260 265 270	816
10	ATG GAC AAC ATT TGT CCT CCT TCA ACG GTT TTC GCT GCC TAC AAT TAC Met Asp Asn Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn Tyr 275 280 285	864
	TAC GCT GGA CCG AAG GAA ATC AGA ATC TAT CCG TAC AAC AAC CAC GAG Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu 290 295 300	864
15	GGA GGA GGC TCT TTC CAA GCG GTT GAA CAG GTG AAA TTC TTG AAA AAA Gly Gly Gly Ser Phe Gln Ala Val Glu Gln Val Lys Phe Leu Lys Lys 305 310 315 320	912
20	CTA TTT GAG AAA GGC TAA Leu Phe Glu Lys Gly 325	930

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 660 NUCLEOTIDES
 (B) TYPE: NUCLEIC ACID
 (C) STRANDEDNESS: SINGLE
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: GENOMIC DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

35

40

45

50

55

EP 0 880 590 B1

	TTG AAG TAC TTC AAA GCC CGG CTT GCC GGC ATC ACC TTG CTC GGC CTG Leu Lys Tyr Phe Lys Ala Arg Leu Ala Gly Ile Thr Leu Leu Gly Leu 1 5 10 15	48
5	CTG GCC TGC ACC TCG GCC TCG CGG CAG ACT GAG CCC ATC GTG TTC GTC Leu Ala Cys Thr Ser Ala Ser Ala Gln Thr Glu Pro Ile Val Phe Val 20 25 30	96
	CAC GGC TAT TCC GGC AGC GCA TCC AAC TGG GAC ACC ATG CTG GGC CGC His Gly Tyr Ser Gly Ser Ala Ser Asn Trp Asp Thr Met Leu Gly Arg 35 40 45	144
10	TTC CCG TCG AAC GGT TAT GCG TCC GGC TCG CTC TAC ACC TTC AAC TAC Phe Arg Ser Asn Gly Tyr Ala Ser Gly Ser Leu Tyr Thr Phe Asn Tyr 50 55 60	192
	AAC TCG TTG GTC AGC AGC AAC CGC ACC AGC GCC AGC GAG CTG CGC AGC Asn Ser Leu Val Ser Ser Asn Arg Thr Ser Ala Ser Glu Leu Arg Ser 65 70 75 80	240
15	TTC GTC AAC ACC GTG CGT TCG CGC CAC GGC AAC GCC CGC ATC GCG CTG Phe Val Asn Thr Val Arg Ser Arg His Gly Asn Ala Arg Ile Ala Leu 85 90 95	288
	GTC GCC CAC TCC AAC GGC GGG CTG GTG TCG CGC TGG TAT CGC GCG GAG Val Ala His Ser Asn Gly Gly Leu Val Ser Arg Trp Tyr Arg Ala Glu 100 105 110	336
20	CTG GGC GGC GAA ACG GCC ACC CGC CGC TTC GTG ACG CTG GGC ACG CCG Leu Gly Gly Glu Thr Ala Thr Arg Arg Phe Val Thr Leu Gly Thr Pro 115 120 125	384
25	CAC CGG GGC ACC ACC TGG GCC TAT GCG TGC TAC AGC CGC CCC GCA TGT TTC His Arg Gly Thr Thr Trp Ala Tyr Ala Cys Tyr Ser Pro Ala Cys Phe 130 135 140	432
30	GAG ATG CGC CCC GGC TCC AGC TTG CTG ACC ACG CTG GGC TCG CGT GCC Glu Met Arg Pro Gly Ser Ser Leu Leu Thr Thr Leu Gly Ser Arg Ala 145 150 155 160	480
	TGC GAC CGC TCG CTG TGG TCG AAC ACC GAC GGC ATC ATC CTG CGG GCG Cys Asp Arg Ser Leu Trp Ser Asn Thr Asp Gly Ile Ile Leu Pro Ala 165 170 175	528
35	TCC AGC GCG CAG TGT GGT GTC AGC ACG CGC ACT GCC GAC GTC AGC CAT Ser Ser Ala Gln Cys Gly Val Ser Thr Arg Thr Ala Asp Val Ser His 180 185 190	576
40	CTC GAC CTG CTG ACC GAC TCT CGC GTG TAC ACG CGT TTG CGC ACG CAG Leu Asp Leu Leu Thr Asp Ser Arg Val Tyr Thr Gln Leu Arg Thr Gln 195 200 205	624
	TTG CAA TGA GGG TGA CGG TGC ACC GAA CGT GCA CCT G Leu Gln End Gly End Arg Cys Thr Glu Arg Ala Pro 210 215 220	661

45 (2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 201 AMINO ACIDS
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

EP 0 880 590 B1

Met Lys Val Lys His Val Ile Val Leu His Gly Leu Tyr Met Ser Gly
1 5 10 15

Leu Val Met Arg Pro Leu Cys Ser Arg Leu Glu Glu Ser Gly Val Lys
5 20 25 30

Val Leu Asn Leu Thr Tyr Asn Thr Arg Asp Pro Asn Arg Asp Ala Ile
10 35 40 45

Phe Thr Gln Ile Asp Glu Phe Ile Ser Asn Glu Pro Ser Ala Leu Val
15 50 55 60

Cys His Ser Met Gly Gly Leu Val Ala Arg Ala Tyr Leu Glu Ala Asn
20 65 70 75 80

Ser Ala Pro Ser His His Val Glu Lys Val Ile Thr Leu Gly Thr Pro
25 85 90 95

His Thr Gly Ser His Ile Ala Glu Lys Met Gln Gln Lys Gly Phe Glu
30 100 105 110

Leu Leu Leu Lys Asn Ser Val Glu Phe Leu Leu Ser Lys Asn Gly Asp
35 115 120 125

Trp Pro Phe Lys Ala Lys Leu Tyr Ser Ile Ala Gly Asp Leu Pro Ile
40 130 135 140

Gly Leu Met Pro Leu Ile Val Lys Gly Ser Arg Ser Asp Gly Thr Val
45 145 150 155 160

Leu Leu Asp Glu Thr Lys Leu Lys Gly Met Ala Glu His Lys Val Phe
50 165 170 175

His Leu Ser His Thr Ser Met Ile Tyr Ser Arg Gln Val Val Asn Tyr
55 180 185 190

Ile Leu Glu Arg Leu Asn Glu Asp Ile
60 195 200

(2) INFORMATION FOR SEQ ID NO:54 :

30

(i) SEQUENCE CHARACTERISTICS

35

- (A) LENGTH: 259 AMINO ACIDS
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

40

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

45

50

55

Met Ile Lys Asn Phe Asp Arg Glu Asn Ser Ser Leu Val Leu Ser Gly
 1 5 10 15
 Gly Gly Ala Leu Gly Ile Ala His Leu Gly Val Leu His Asp Leu Glu
 20 25 30
 Lys Gln Asn Ile Val Pro Asn Glu Ile Val Gly Thr Ser Met Gly Gly
 35 40 45
 Ile Ile Gly Ala Ser Met Ala Ile Gly Met Lys Glu Lys Glu Ile Leu
 50 55 60
 10 Glu Glu Ile Lys Asn Phe Ser Asn Val Phe Asn Trp Ile Lys Phe Ser
 65 70 75 80
 Phe Ser Gly Asn Ser Val Val Asp Asn Glu Lys Ile Ala Lys Ile Phe
 85 90 95
 15 Asp Thr Leu Phe Lys Asp Arg Lys Met Thr Asp Thr Val Ile Pro Leu
 100 105 110
 Lys Leu Ile Ala Thr Asn Leu His Asn Gly His Lys Lys Val Phe Thr
 115 120 125
 Ala Ser Asp Asp Val Leu Ile Lys Asp Ala Ile Leu Ser Thr Met Ala
 130 135 140
 20 Ile Pro Gly Val Phe Glu Glu His Ile Ile Asp Gly Glu Thr Tyr Gly
 145 150 155 160
 Asp Gly Phe Leu Cys Glu Asn Leu Gly Val Asn Glu Ala Thr Phe Asn
 165 170 175
 25 Asp Val Leu Ala Val Asp Val Met Gly Glu Asn Ser Phe Glu Lys Ala
 180 185 190
 Met Pro Asp Asn Phe Phe Lys Thr Ser Asn Val Leu Glu Met Phe Glu
 195 200 205
 Lys Ser Met Arg Leu Phe Ile Tyr Asn Gln Thr Gln Thr His Ile Lys
 210 215 220
 30 Asn Ala Asn Lys Asn Ile Tyr Leu Ile Glu Pro Val Thr Lys Glu Tyr
 225 230 235 240
 Lys Thr Phe Gln Phe His Lys Lys Glu Ile Arg Ala Leu Gly Leu
 245 250 255
 Gly Leu Leu
 35

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 301 AMINO ACIDS
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

40 Met Pro Leu His Pro Lys Val Lys Lys Leu Leu Ser Gln Leu Pro Pro
 1 5 10 15
 Gln Asp Phe Ser Arg Asn Val Gln Asp Leu Arg Lys Ala Trp Asp Leu
 20 25 30
 Pro Phe Ser Gly Arg Arg Glu Thr Leu Lys Arg Val Glu Asp Leu Glu
 35 40 45
 50 Ile Pro Thr Arg Asp Ala Arg Ile Arg Ala Arg Val Tyr Thr Pro Ser
 55 60

EP 0 880 590 B1

Ser Lys Glu Asn Leu Pro Val Leu Val Tyr Tyr His Gly Gly Gly Phe
 65 70 75 80

5 Val Phe Gly Ser Val Asp Ser Tyr Asp Gly Leu Ala Ser Leu Ile Ala
 85 90 95

Lys Glu Ser Gly Ile Ala Val Ile Ser Val Glu Tyr Arg Leu Ala Pro
 100 105 110

10 Glu His Lys Phe Pro Thr Ala Val Asn Asp Ser Trp Asp Ala Leu Leu
 115 120 125

Trp Ile Ala Glu Asn Gly Gly Lys Leu Gly Leu Asp Thr Ser Arg Leu
 130 135 140

Ala Val Ala Gly Asp Ser Ala Gly Gly Asn Leu Ser Ala Val Val Ser
 145 150 155 160

15 Leu Leu Asp Arg Asp Gln Gly Lys Gly Leu Val Ser Tyr Gln Val Leu
 165 170 175

Ile Tyr Pro Ala Val Asn Met Val Asp Asn Ser Pro Ser Val Arg Glu
 180 185 190

Tyr Gly Glu Gly Tyr Phe Leu Thr Arg Ser Met Met Asn Trp Phe Gly
 195 200 205

20 Thr Met Tyr Phe Ser Ser Gly Arg Glu Ala Val Ser Pro Tyr Ala Ser
 210 215 220

Pro Ala Leu Ala Asp Leu His Asn Leu Pro Pro Ser Leu Val Ile Thr
 225 230 235 240

25 Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Thr Tyr Ser His Ser
 250 255 260

Leu Asn Glu Ala Gly Asn Val Ser Thr Leu Val Arg Tyr Gln Gly Met
 265 270 275

30 Ile His Gly Phe Leu Ser Phe Tyr Glu Trp Ile Thr Ala Gly Lys Leu
 280 285 290

Ala Ile His His Ile Ala Gly Val Leu Arg Ser Val Leu
 295 300 305

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 326 AMINO ACIDS
 - (B) TYPE: AMINO ACID
 - (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Val Ala Phe Phe Asp Met Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro
 1 5 10 15

Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Arg Glu Thr Leu
 20 25 30

5 Lys Glu Ser Glu Gly Phe Pro Leu Asp Pro Val Phe Glu Lys Val Asp
 35 40 45

Phe His Leu Lys Thr Val Glu Thr Tyr Asp Val Thr Phe Ser Gly Tyr
 50 55 60

10 Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Ala Glu
 65 70 75 80

Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg
 85 90 95

Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys
 100 105 110

15

20 Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Met Lys Gly Asp
 115 120 125

Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe
 130 135 140

25 Met Thr Arg Gly Ile Leu Asp Pro Gly Thr Tyr Tyr Arg Arg Val
 145 150 155 160

Phe Val Asp Ala Val Arg Ala Val Glu Ala Ala Ile Ser Phe Pro Arg
 165 170 175

Val Asp Ser Arg Lys Val Val Ala Gly Gly Ser Gln Gly Gly Gly
 180 185 190

30 Ile Pro Leu Ala Val Ser Ala Leu Ser Asn Arg Val Lys Ala Leu Leu
 195 200 205

Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val
 210 215 220

35 Asp Thr His Pro Tyr Val Glu Ile Thr Asn Phe Leu Lys Thr His Arg
 225 230 235 240

Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val
 245 250 255

Asn Phe Ala Ala Arg Ala Lys Val Pro Ala Leu Phe Ser Val Gly Leu
 260 265 270

40 Met Asp Thr Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn His
 275 280 285

Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu
 290 295 300

45 Gly Gly Gly Ser Phe Gln Ala Ile Glu Gln Val Lys Phe Leu Lys Arg
 305 310 315 320

Leu Phe Glu Glu Gly
 325

(2) INFORMATION FOR SEQ ID NO:57:

- 50 (i) SEQUENCE CHARACTERISTICS
- (A) LENGTH: 298 AMINO ACIDS
 - (B) TYPE: AMINO ACID
 - 55 (D) TOPOLOGY: LINEAR

- (ii) MOLECULE TYPE: PROTEIN

EP 0 880 590 B1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

5 Met Arg Thr Leu Ser Phe Gly Pro Met Thr Thr Gly Gly Ser Ile His
1 5 10 15
Met Ala Thr Met Asp Val Met Arg Gly Pro Gly Met Gln Arg Leu Ser
20 25 30
Gln Gly Ala Arg Glu Ala Ala Asn His Pro Trp Ala Lys Arg Leu Gly
35 40 45
10 Arg Met Gly Tyr Ala Ala Lys Gly Ala Val Tyr Ala Ile Ile Gly Val
50 55 60
Leu Ala Leu Lys Leu Ala Gly Glu Gly Gly Arg Thr Thr Asp Ser
65 70 75 80
His Gly Ala Val Asn Thr Val Ala His Gly Pro Phe Gly Val Ala Leu
85 90 95
15 Leu Ala Val Leu Val Val Gly Leu Leu Gly Tyr Val Val Trp Arg Phe
100 105 110
Ala Gln Ala Phe Val Asp Thr Glu Asp Lys Gly Ser Asp Ala Lys Gly
115 120 125
20

Ile Ala Thr Arg Ala Met Tyr Phe Leu Ser Gly Cys Ile Tyr Ala Ser
130 135 140
25 Leu Ala Phe Phe Ala Ala Gln Ser Leu Val Gly Ala Ala His Gly Arg
145 150 155 160
Ser Lys Gly Thr Gln Gly Trp Thr Ala Thr Leu Met Glu Gln Pro Phe
165 170 175
30 Gly Arg Val Leu Val Ala Leu Val Gly Leu Gly Ile Val Gly Phe Ala
180 185 190
Leu Lys Gln Phe His Thr Ala Trp Lys Ala Lys Phe Arg Glu Lys Leu
195 200 205
Thr Leu Thr Gly Leu Ala Ala Arg Lys Gln His His Ile Glu Arg Met
210 215 220
35 Cys Gln Phe Gly Ile Ala Ala Arg Gly Val Val Phe Ala Val Ile Gly
225 230 235 240
Gly Phe Leu Val Arg Ser Ala Val Asp Ala Asn Pro Gly Glu Ala Lys
245 250 255
40 Gly Leu Gly Glu Ala Leu Ala Val Val Ala Arg Gln Pro Ser Gly Asp
260 265 270
Val Leu Leu Gly Val Val Ala Ala Gly Ile Val Ala Tyr Ala Ala Tyr
275 280 285
45 Leu Phe Leu Gln Ala Arg Tyr Arg Glu Leu
290 295

(2) INFORMATION FOR SEQ ID NO:58:

50 (i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 304 AMINO ACIDS
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

55 (ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

EP 0 880 590 B1

	Met Ser Lys Phe Ala Ile Leu Trp Ala Leu Ile Thr Ala Tyr Leu Pro
1	5 10 15
	Glu Pro Val Met Lys Leu Val Tyr Leu Gly Arg Arg Glu Thr Leu Gly
5	20 25 30
	Ala Arg Thr Leu Asp Val Lys Ala Gln Ala Val Gly Arg Leu Ala Asn
	35 40 45
	Ala Thr Arg Pro Val Gly Val Ile Pro Thr Val Glu Glu Ser Arg Lys
	50 55 60
10	Met Thr Asp Lys Ala Val Ser Leu Phe Asp Gln Pro Ala Pro Glu Leu
	65 70 75 80
	Phe Arg Lys Lys Asp Ile Gln Ile Asp Gly Ala Glu Gly Pro Ile Asp
	85 90 95
	Ala Arg Ile Tyr Ser Gly Pro Ala Lys His Arg Pro Arg Pro Ile Leu
15	100 105 110
	Val Tyr Phe His Gly Gly Gly Trp Val Gln Gly Asn Leu Asp Ser His
	115 120 125
	Asp Gly Val Cys Gly Lys Leu Ala Lys Trp Ala Asn Cys Ile Val Ile
	130 135 140
20	Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Cys Ala Pro
	145 150 155 160
	Leu Asp Ala Ile Ala Ala Tyr Lys Trp Val Arg Ala Asn Ala Thr Asn
	165 170 175
25	Leu Gly Gly Asp Pro Glu Arg Ile Gly Val Gly Gly Asp Ser Ala Gly
	180 185 190
	Gly Asn Leu Ala Ala Val Val Cys Gln Gln Thr Ala Met Asn Gly Glu
30	195 200 205
	Arg Thr Pro Asp Leu Gln Val Leu Ile Tyr Pro Ala Leu Asp Ala Arg
	210 215 220
	Met Ile Ser Thr Ser Met Glu Glu Leu Arg Asp Ala Tyr Ile Leu Pro
	225 230 235 240
35	Lys Ser Arg Met Glu Tyr Phe Leu Gly Leu Tyr Thr Arg Gly Pro Asp
	245 250 255
	Asp Ile Glu Asp Leu Arg Met Ser Pro Ile Leu Arg Asp Thr Val Ala
	260 265 270
40	Asp Gln Pro Gln Ala Cys Ile Val Thr Cys Gly Phe Asp Pro Ala Arg
	275 280 285
	Arg Arg Glu His Leu Arg Arg Thr Leu Asn Cys Arg Gly Asp Arg Arg
	290 295 300

45 (2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 308 AMINO ACIDS
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

EP 0 880 590 B1

	Val Ser Ile Arg Leu Arg Leu Leu Asn Trp Phe Leu Asn Thr Phe Glu			
1	5	10	15	
	Lys Pro Lys Leu Ala Ala Ala Lys Thr Pro Asp Asp Leu Arg Lys Ser			
5	20	25	30	
	Phe Glu Leu Lys Ala Arg Phe Leu Phe Pro Ala Pro Arg Lys Thr Arg			
	35	40	45	
	Phe Ser His Asp Val Leu Gln Ser Gly Ile Gly Ser Val Asn Ala Gln			
10	50	55	60	
	Trp Ala Lys Ser Lys Ser Ala Ser Asp Asp Arg Val Ile Leu Tyr Phe			
	65	70	75	80
	His Gly Gly Tyr Val Phe Gly Ser Pro Lys Thr His Arg Ala Met			
15	85	90	95	
	Leu Ala Arg Leu Ser Ala Met Thr Gly Leu Ser Ala Cys Leu Pro Asp			
	100	105	110	
	Tyr Arg Leu Ala Pro Glu His Pro Phe Pro Ala Ala Ile Glu Asp Ala			
	115	120	125	
	Val Leu Ser Tyr Lys Cys Leu Leu Glu Arg Ala Ile Glu Pro Gln Asn			
20	130	135	140	
	Ile Ile Leu Gly Gly Asp Ser Ala Gly Gly Leu Val Leu Ala Leu			
	145	150	155	160
	Leu Ala Glu Ile Lys Ala Gln Ser Leu Pro Lys Pro Ala Gly Val Phe			
	165	170	175	
	Ala Leu Ser Pro Leu Val Asp Leu Ser Phe Ser Gly Leu Ser Phe Ser			
25	180	185	190	
	Lys Asn Ala Gln Thr Asp Val Met Leu Pro Ala Ser Arg Ala Ala Asp			
	195	200	205	
	Met Ala Thr Leu Tyr Leu Asp Gly Ala Asp Ala Asp Pro Arg Ala			
	210	215	220	
30				
	Ser Pro Leu Gln Ala Asp Phe Ser Gly Met Pro Pro Val Phe Leu Thr			
	225	230	235	240
35				
	Ala Ser Asp Ser Glu Ile Leu Leu Asp Asp Cys Leu Arg Met Ala Asp			
	245	250	255	
	His Leu Arg Ala Gln Gly Val Val Val Thr Asp Arg Ile Val Glu Asn			
	260	265	270	
40				
	His Pro His Val Trp His Ile Phe Gln Arg Leu Leu Pro Glu Ala Asp			
	275	280	285	
	Gln Gly Leu Arg Ala Ile Ala Ala Trp Ile Lys Pro Leu Leu Ser Gly			
	290	295	300	
	Ser Asn Glu Ser			
45	305			
	(2) INFORMATION FOR SEQ ID NO:60:			
	(i) SEQUENCE CHARACTERISTICS			
50	(A) LENGTH: 237 AMINO ACIDS			
	(B) TYPE: AMINO ACID			
	(D) TOPOLOGY: LINEAR			
55	(ii) MOLECULE TYPE: PROTEIN			
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:			

Met Leu Thr Phe Asn Val Leu Tyr Gly Met Met Lys Gln Lys Leu Ala
 1 5 10 15

Ala Ile Leu Met Phe Leu Gly Leu Ser Ala Ala Glu Ala Gln Asp Trp
 20 25 30

5 Pro Asp Leu Gln Lys Tyr Arg Ser Ala Asn Lys Glu Ala Lys Leu Leu
 35 40 45

Pro Lys Glu Asn Arg Lys Val Val Phe Met Gly Asn Ser Ile Thr Glu
 50 55 60

10 Ala Trp Ile Ser Gln Arg Pro Glu Phe Phe Ser Glu Asn Gly Phe Ile
 65 70 75 80

Gly Arg Gly Ile Ser Gly Gln Thr Thr Pro Gln Met Leu Leu Arg Phe
 85 90 95

15 Arg Gln Asp Val Ile Asp Leu Gln Pro Lys Ala Val Val Ile Leu Ala
 100 105 110

Gly Thr Asn Asp Val Ala Gln Asn Thr Gly Pro Met Thr Ile Glu Glu
 115 120 125

Ser Leu Ala Asn Ile Lys Ser Met Val Glu Leu Ala Gln Ala Asn Gly
 130 135 140

20 Ile Thr Pro Val Leu Cys Thr Val Leu Pro Ala Asp Arg Phe Ser Trp
 145 150 155 160

Arg Pro Glu Leu Thr Pro Ala Glu Thr Ile Ile Ala Leu Asn Gln Leu
 165 170 175

25 Ile Lys Gln Tyr Ala Glu Ala Gln Gly Leu Ala Leu Val Asp Tyr His
 180 185 190

Ala Ala Leu Thr Asn Lys Gly Gly Leu Pro Val Lys Tyr Gly Glu
 195 200 205

Asp Gly Val His Pro Asn Val Ala Gly Tyr Gln Val Met Glu Asn Ile
 210 215 220

30 Val Leu Pro Val Ile Ser Ser Glu Leu Ala Lys Leu Lys
 225 230 235

(2) INFORMATION FOR SEQ ID NO:61:

35

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS

40

- (A) LENGTH: 326 AMINO ACIDS
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

45

(ii) MOLECULE TYPE: PROTEIN

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

55

EP 0 880 590 B1

	Met Ala Phe Phe Asp Leu Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro
1	5 10 15
5	Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Glu Glu Thr Leu
	20 25 30
	Ala Glu Ser Glu Lys Phe Pro Leu Asp Pro Val Phe Glu Arg Met Glu
	35 40 45
	Ser His Leu Lys Thr Val Glu Ala Tyr Asp Val Thr Phe Ser Gly Tyr
	50 55 60
10	Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Glu Glu
	65 70 75 80
	Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg
	85 90 95
15	Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys
	100 105 110
	Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Leu Lys Gly Asp
	115 120 125
	Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe
	130 135 140
20	Met Thr Arg Gly Ile Leu Asp Pro Arg Thr Tyr Tyr Tyr Arg Arg Val
	145 150 155 160
	Phe Thr Asp Ala Val Arg Ala Val Glu Ala Ala Ala Ser Phe Pro Gln
	165 170 175
25	Val Asp Gln Glu Arg Ile Val Ile Ala Gly Gly Ser Gln Gly Gly Gly
	180 185 190
	Ile Ala Leu Ala Val Ser Ala Leu Ser Lys Lys Ala Lys Ala Leu Leu
	195 200 205
	Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val
	210 215 220
30	Asp Thr His Pro Tyr Ala Glu Ile Thr Asn Phe Leu Lys Thr His Arg
	225 230 235 240
	Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val
	245 250 255
35	Asn Phe Ala Ala Arg Ala Lys Ile Pro Ala Leu Phe Ser Val Gly Leu
	260 265 270
	Met Asp Asn Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn Tyr
	275 280 285
	Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu
	290 295 300
40	Gly Gly Gly Ser Phe Gln Ala Val Glu Gln Val Lys Phe Leu Lys Lys
	305 310 315 320
	Leu Phe Glu Lys Gly
	325

45

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS

50

(A) LENGTH: 220 AMINO ACIDS

(B) TYPE: AMINO ACID

55

(D) TOPOLOGY: LINEAR

EP 0 880 590 B1

(iii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

5 Leu Lys Tyr Phe Lys Ala Arg Leu Ala Gly Ile Thr Leu Leu Gly Leu
 1 5 10 15

Leu Ala Cys Thr Ser Ala Ser Ala Gln Thr Glu Pro Ile Val Phe Val
 20 25 30

His Gly Tyr Ser Gly Ser Ala Ser Asn Trp Asp Thr Met Leu Gly Arg
 35 40 45

10 Phe Arg Ser Asn Gly Tyr Ala Ser Gly Ser Leu Tyr Thr Phe Asn Tyr
 50 55 60

Asn Ser Leu Val Ser Ser Asn Arg Thr Ser Ala Ser Glu Leu Arg Ser
 65 70 75 80

Phe Val Asn Thr Val Arg Ser Arg His Gly Asn Ala Arg Ile Ala Leu
 85 90 95

Val Ala His Ser Asn Gly Gly Leu Val Ser Arg Trp Tyr Arg Ala Glu
 100 105 110

20 Leu Gly Gly Glu Thr Ala Thr Arg Arg Phe Val Thr Leu Gly Thr Pro
 115 120 125

His Arg Gly Thr Thr Trp Ala Tyr Ala Cys Tyr Ser Pro Ala Cys Phe
 130 135 140

Glu Met Arg Pro Gly Ser Ser Leu Leu Thr Thr Leu Gly Ser Arg Ala
 145 150 155 160

25 Cys Asp Arg Ser Leu Trp Ser Asn Thr Asp Gly Ile Ile Leu Pro Ala
 165 170 175

Ser Ser Ala Gln Cys Gly Val Ser Thr Arg Thr Ala Asp Val Ser His
 180 185 190

Leu Asp Leu Leu Thr Asp Ser Arg Val Tyr Thr Gln Leu Arg Thr Gln
 195 200 205

30 Leu Gln End Gly End Arg Cys Thr Glu Arg Ala Pro G
 210 215 220

35 SEQUENCE LISTING

[0133]

40 (1) GENERAL INFORMATION:

(i) APPLICANT: Diversa Corporation

45 (ii) TITLE OF INVENTION: ESTERASES

(iii) NUMBER OF SEQUENCES: 62

(iv) CORRESPONDENCE ADDRESS:

50 (A) ADDRESSEE: Morrison & Foerster LLP
 (B) STREET: 3811 Valley Centre Drive, Suite 500
 (C) CITY: San Diego
 (D) STATE: CA
 (E) COUNTRY: USA
55 (F) ZIP: 92130

(v) COMPUTER READABLE FORM:

EP 0 880 590 B1

- (A) MEDIUM TYPE: Diskette
- (B) COMPUTER: IBM Compatible
- (C) OPERATING SYSTEM: Windows 95
- (D) SOFTWARE: FastSEQ for Windows Version 2.0b

5

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: US97/02039
- (B) FILING DATE: 11-FEB-1997

10

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: 08/602,359
- (B) FILING DATE: 16-FEB-1996

15

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Einhorn, Ph.D., D.D.S., Gregory
- (B) REGISTRATION NUMBER: 38,440
- (C) REFERENCE/DOCKET NUMBER: 09010/010WO1

20

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 858 720 5133
- (B) TELEFAX: 858 720 5125

25

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

30

- (A) LENGTH: 52 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

40

CCGAGAATTTC ATTAAAGAGG AGAAATTAAC TATGTCTTTA AACAAAGCACT CT

52

(2) INFORMATION FOR SEQ ID NO:2:

45

(i) SEQUENCE CHARACTERISTICS:

50

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CGGAAGATCT CTATCGTTTA GTGTATGATT T

31

EP 0 880 590 B1

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 52 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

15 CCGAGAATTC ATAAAGAGG AGAAATTAAAC TATGAAACTC CTTGAGCCCA CA

52

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 31 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

30 CGGAAGATCT CGCCGGTACA CCATCAGCCA C

31

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 52 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

45 CCGAGAATTC ATAAAGAGG AGAAATTAAAC TATGCCATAT GTAGGAATG GT

52

50 (2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 53 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

EP 0 880 590 B1

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

5 CGGAGGTACC TTAGAACTGT GCTGAAGAAA TAAATTCGTC CATTGCTCTA TTA

53

(2) INFORMATION FOR SEQ ID NO:7:

10 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 49 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: cDNA

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

25 CGGAGGTACC TTAGAACTGT GCTGAAGAAA TAAATTCGTC CATTGCTCT

49

(2) INFORMATION FOR SEQ ID NO:8:

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: cDNA

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

40 CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGAGATTG AGGAAATTG AAG

53

(2) INFORMATION FOR SEQ ID NO:9:

45 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

55 CGGAGGTACC CTATTCAGAA AGTACCTCTA A

31

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

EP 0 880 590 B1

- (A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

10

CCGAGAATTC ATTAAGAGG AGAAATTAAC TATGTTAAC ATCAATGTCT TT

52

(2) INFORMATION FOR SEQ ID NO:11:

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

25

CGGAAGATCT TTAAGGATTG TCCCTGGGTA G

31

(2) INFORMATION FOR SEQ ID NO:12:

30

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: cDNA

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCGAGAATTC ATTAAGAGG AGAAATTAAC TATGGAGGTT TACAAGGCCA AA

52

45

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

50

- (A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

55

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGAGGTACC TTATTGAGCC GAAGAGTACG A

31

(2) INFORMATION FOR SEQ ID NO:14:

5

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CCGAGAATTC ATAAAGAGG AGAAATTAAC TATGATTGGC AATTTGAAAT TGA

53

20

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CGGAGGTACC TTAAAGTGCT CTCATATCCC C

31

35

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

40

- (A) LENGTH: 52 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

45

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

50

CCGAGAATTC ATAAAGAGG AGAAATTAAC TATGCCAGCT AATGACTCAC CC

52

(2) INFORMATION FOR SEQ ID NO:17:

55

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

EP 0 880 590 B1

(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGGAAGATCT TCAACAGGCT CCAAATAATT TC

32

10

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

15

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

25

CGGAAGATCT ACAGGCTCCA AATAATTTC

29

30

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

35

- (A) LENGTH: 52 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

40

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

45

CCGAGAACCTT ATTAAAGAGG AGAAATTAAC TATGCTTGAT ATGCCAATCG AC

52

50

(2) INFORMATION FOR SEQ ID NO:20:

55

(i) SEQUENCE CHARACTERISTICS:

60

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

65

(ii) MOLECULE TYPE: cDNA

70

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGGAGGTACC CTAGTCGAAC AGAAGAAGAG C

31

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

15 CCGAGAAATTTC ATTAAAGAGG AGAAAATTAAC TATGCCCTA GATCCTAGAA TT

52

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

30 CGGAGGTACC TTAAATTTA TCATAAAATA C

31

(2) INFORMATION FOR SEQ ID NO:23:

35 (i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 555 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

45 (ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...552

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

55

EP 0 880 590 B1

ATG TCT TTA AAC AAG CAC TCT TGG ATG GAT ATG ATA ATA TTT ATT CTC Met Ser Leu Asn Lys His Ser Trp Met Asp Met Ile Ile Phe Ile Leu 1 5 10 15	48
5 AGC TTT TCT TTC CCA TTA ACA ATG ATC GCA TTA GCT ATC TCT ATG TCG Ser Phe Ser Phe Pro Leu Thr Met Ile Ala Leu Ala Ile Ser Met Ser 20 25 30	96
10 TCA TGG TTT AAT ATA TGG AAT AAT GCA TTA AGC GAT CTA GGA CAT GCT Ser Trp Phe Asn Ile Trp Asn Asn Ala Leu Ser Asp Leu Gly His Ala 35 40 45	144
15 GTT AAA AGC AGT GTT GCT CCA ATA TTC AAT CTA GGT CTT GCA ATT GGT Val Lys Ser Ser Val Ala Pro Ile Phe Asn Leu Gly Leu Ala Ile Gly 50 55 60	192
20 GGG ATA CTA ATT GTT ATA GTT GGT TTA AGA AAT CTT TAT TCG TGG AGT Gly Ile Leu Ile Val Ile Val Gly Leu Arg Asn Leu Tyr Ser Trp Ser 65 70 75 80	240
AGA GTT AAA GGA TCT TTA ATC ATA TCC ATG GGT GTA TTT CTT AAC TTA Arg Val Lys Gly Ser Leu Ile Ile Ser Met Gly Val Phe Leu Asn Leu 85 90 95	288
25 ATA GGG GTT TTC GAC GAA GTA TAT GGT TGG ATA CAT TTC CTA GTC TCA Ile Gly Val Phe Asp Glu Val Tyr Gly Trp Ile His Phe Leu Val Ser 100 105 110	336
30 GTA TTG TTT TTC TTA TCA ATA ATA GCA TAT TTC ATA GCT ATA TCA ATA Val Leu Phe Phe Leu Ser Ile Ile Ala Tyr Phe Ile Ala Ile Ser Ile 115 120 125	384
35 CTT GAC AAA TCA TGG ATA GCT GTT CTA CTA ATA ATA GGT CAT ATT GCA Leu Asp Lys Ser Trp Ile Ala Val Leu Leu Ile Ile Gly His Ile Ala 130 135 140	432
40 ATG TGG TAT CTA CAC TTT GCT TCA GAG ATT CCG AGA GGT GCG GCT ATT Met Trp Tyr Leu His Phe Ala Ser Glu Ile Pro Arg Gly Ala Ala Ile 145 150 155 160	480
CCC GAG TTA TTA GCG GTA TTC TCG TTT TTA CCA TTC TAT ATA AGA GAC Pro Glu Leu Leu Ala Val Phe Ser Phe Leu Pro Phe Tyr Ile Arg Asp 165 170 175	528
45 TAT TTT AAA TCA TAC ACT AAA CGA TAG Tyr Phe Lys Ser Tyr Thr Lys Arg 180	555

(2) INFORMATION FOR SEQ ID NO:24:

- 50 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1041 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- 55 (ii) MOLECULE TYPE: Genomic DNA

EP 0 880 590 B1

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...1038

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

10

ATG AAA CTC CTT GAG CCC ACA AAT ACC TCC TAC ACG CTG TTA CAG GAT
Met Lys Leu Leu Glu Pro Thr Asn Thr Ser Tyr Thr Leu Leu Gln Asp
1 5 10 15

48

15

TTA GCA TTG CAT TTT GCA TTT TAC TGG TTT CTG GCC GTG TAT ACG TGG

96

20

25

30

35

40

45

50

55

EP 0 880 590 B1

	Leu Ala Leu His Phe Ala Phe Tyr Trp Phe Leu Ala Val Tyr Thr Trp		
	20	25	30
5	TTA CCC GGT GTC CTA GTC CGG GGC GTA GCT GTG GAC ACA GGG GTG GCT		144
	Leu Pro Gly Val Leu Val Arg Gly Val Ala Val Asp Thr Gly Val Ala		
	35	40	45
10	CGG GTG CCT GGG CTC GGC CGG CGC GGT AAG AGG CTG CTC CTG GCC GCT		192
	Arg Val Pro Gly Leu Gly Arg Arg Gly Lys Arg Leu Leu Ala Ala		
	50	55	60
15	GTG GCT GTC TTG GCG CTT GTT GTG TCC GTT GTC CCG GCT TAT GTG		240
	Val Ala Val Leu Ala Leu Val Val Ser Val Val Val Pro Ala Tyr Val		
	65	70	75
	80		
20	GCG TAT AGT AGT CTG CAC CCG GAG AGC TGT CGG CCC GTT GCG CCG GAG		288
	Ala Tyr Ser Ser Leu His Pro Glu Ser Cys Arg Pro Val Ala Pro Glu		
	85	90	95
25	GGG CTC ACC TAC AAA GAG TTC AGC GTG ACC GCG GAG GAT GGC TTG GTG		336
	Gly Leu Thr Tyr Lys Glu Phe Ser Val Thr Ala Glu Asp Gly Leu Val		
	100	105	110
30	GTT CGG GGC TGG GTG CTG GGC CCC GGC GCT GGG GGC AAC CCG GTG TTC		384
	Val Arg Gly Trp Val Leu Gly Pro Gly Ala Gly Gly Asn Pro Val Phe		
	115	120	125
35	GTT TTG ATG CAC GGG TAT ACT GGG TGC CGC TCG GCG CCC TAC ATG GCT		432
	Val Leu Met His Gly Tyr Thr Gly Cys Arg Ser Ala Pro Tyr Met Ala		
	130	135	140
40	GTG CTG GCC CGG GAG CTC GTG GAG TGG GGG TAC CCG GTG GTT GTG TTC		480
	Val Leu Ala Arg Glu Leu Val Glu Trp Gly Tyr Pro Val Val Val Phe		
	145	150	155
	160		
45	GAC TTC CGG GGC CAC GGG GAG AGC GGG GGC TCG ACG ACG ATT GGG CCC		528
	Asp Phe Arg Gly His Gly Glu Ser Gly Gly Ser Thr Thr Ile Gly Pro		
	165	170	175
50	CGG GAG GTG CTG GAT GCC CGG GCT GTG GTG GGC TAT GTC TCG GAG CGG		576
	Arg Glu Val Leu Asp Ala Arg Ala Val Val Gly Tyr Val Ser Glu Arg		
	180	185	190
55	TTC CCC GGC CGC CGG ATA ATA TTG GTG GGG TTC AGT ATG GGC GGC GCT		624
	Phe Pro Gly Arg Arg Ile Ile Leu Val Gly Phe Ser Met Gly Gly Ala		
	195	200	205
60	GTA GCG ATC GTG GAG GGT GCT GGG GAC CCG CGG GTC TAC GCG GTG GCT		672
	Val Ala Ile Val Glu Gly Ala Gly Asp Pro Arg Val Tyr Ala Val Ala		
	210	215	220
65	GCT GAT AGC CCG TAC TAT AGG CTC CGG GAC GTC ATA CCC CGG TGG CTG		720
	Ala Asp Ser Pro Tyr Tyr Arg Leu Arg Asp Val Ile Pro Arg Trp Leu		
	225	230	235
			240

EP 0 880 590 B1

	GAG TAC AAG ACG CCG CTG CCG GGC TGG GTG GGT GTG CTG GCC GGG TTC Glu Tyr Lys Thr Pro Leu Pro Gly Trp Val Gly Val Leu Ala Gly-Phe 245 250 255	768
5	TAC GGG AGG CTG ATG GCG GGC GTT GAC CTC GGC TTC GGC CCC GCT GGG Tyr Gly Arg Leu Met Ala Gly Val Asp Leu Gly Phe Gly Pro Ala Gly 260 265 270	816
10	GTG GAG CGC GTG GAT AAG CCG TTG CTG GTG GTG TAT GGG CCC CGG GAC Val Glu Arg Val Asp Lys Pro Leu Leu Val Val Tyr Gly Pro Arg Asp 275 280 285	864
15	CCG CTG GTG ACG CGG GAC GAG GCG AGG AGC CTG GCG TCC CGT AGC CCG Pro Leu Val Thr Arg Asp Glu Ala Arg Ser Leu Ala Ser Arg Ser Pro 290 295 300	912
20	TGT GGC CGT CTC GTC GAG GTT CCT GGG GCT GGC CAC GTG GAG GCC GTG Cys Gly Arg Leu Val Glu Val Pro Gly Ala Gly His Val Glu Ala Val 305 310 315 320	960
	GAT GTG CTC GGG CCG GGC CGC TAC GCA GAC ATG CTG ATA GAG CTG GCG Asp Val Leu Gly Pro Gly Arg Tyr Ala Asp Met Leu Ile Glu Leu Ala 325 330 335	1008
25	CAC GAG GAG TGC CCT CCG GGG GCC GGT GGC TGA His Glu Glu Cys Pro Pro Gly Ala Gly Gly 340 345	1041

30 (2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 789 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

40 (ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...786

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

	ATG CCA TAT GTT AGG AAT GGT GGT GTA AAT ATC TAT TAT GAA CTG GTG Met Pro Tyr Val Arg Asn Gly Gly Val Asn Ile Tyr Tyr Glu Leu Val 1 5 10 15	48
50	GAT GGA CCT GAG CCA CCA ATT GTC TTT GTT CAC GGA TGG ACA GCA AAT Asp Gly Pro Glu Pro Pro Ile Val Phe Val His Gly Trp Thr Ala Asn 20 25 30	96
55	ATG AAT TTT TGG AAA GAG CAA AGA CGT TAT TTT GCA GGC AGG AAT ATG Met Asn Phe Trp Lys Glu Gln Arg Arg Tyr Phe Ala Gly Arg Asn Met	144

EP 0 880 590 B1

	35	40	45	
5	ATG TTG TTT GTC GAT AAC AGA GGT CAT GGC AGG TCC GAT AAG CCA CTT Met Leu Phe Val Asp Asn Arg Gly His Gly Arg Ser Asp Lys Pro Leu 50 55 60			192
10	GGA TAC GAT TTC TAC AGA TTT GAG AAC TTC ATT TCA GAT TTA GAT GCG Gly Tyr Asp Phe Tyr Arg Phe Glu Asn Phe Ile Ser Asp Leu Asp Ala 65 70 75 80			240
15	GTT GTT AGG GAG ACT GGA GTG GAG AAA TTT GTT CTC GTC GGA CAT TCA Val Val Arg Glu Thr Gly Val Glu Lys Phe Val Leu Val Gly His Ser 85 90 95			288
20	TTC GGA ACA ATG ATC TCT ATG AAG TAC TGT TCG GAG TAT CGG AAT CGG Phe Gly Thr Met Ile Ser Met Lys Tyr Cys Ser Glu Tyr Arg Asn Arg 100 105 110			336
25	GTT CTT GCT CTA ATC CTC ATA GGT GGT GGG AGC AGA ATA AAG CTT CTA Val Leu Ala Leu Ile Leu Ile Gly Gly Ser Arg Ile Lys Leu Leu 115 120 125			384
30	CAC AGA ATT GGA TAT CCT TTA GCA AAG ATT CTT GCA TCC ATT GCA TAC His Arg Ile Gly Tyr Pro Leu Ala Lys Ile Leu Ala Ser Ile Ala Tyr 130 135 140			432
35	AAG AAG TCT TCA AGA TTG GTC GCA GAT CTT TCC TTT GGC AAA AAT GCT Lys Lys Ser Ser Arg Leu Val Ala Asp Leu Ser Phe Gly Lys Asn Ala 145 150 155 160			480
40	GGT GAA CTT AAA GAG TGG GGA TGG AAA CAG GCA ATG GAT TAT ACA CCC Gly Glu Leu Lys Glu Trp Gly Trp Lys Gln Ala Met Asp Tyr Thr Pro 165 170 175			528
45	TCC TAC GTG GCA ATG TAC ACG TAC AGA ACT CTA ACG AAA GTG AAT CTT Ser Tyr Val Ala Met Tyr Thr Tyr Arg Thr Leu Thr Lys Val Asn Leu 180 185 190			576
50	GAA AAT ATC TTG GAG AAA ATA GAC TGT CCA ACA CTG ATT ATC GTT GGA Glu Asn Ile Leu Glu Lys Ile Asp Cys Pro Thr Leu Ile Ile Val Gly 195 200 205			624
55	GAA GAG GAT GCA CTA TTG CCC GTT AGC AAA TCA GTT GAG CTG AGC AGG Glu Glu Asp Ala Leu Leu Pro Val Ser Lys Ser Val Glu Leu Ser Arg 210 215 220			672
	AGG ATA GAA AAC TCA AAG CTT GTG ATC ATC CCA AAC TCG GGG CAT TGC Arg Ile Glu Asn Ser Lys Leu Val Ile Ile Pro Asn Ser Gly His Cys 225 230 235 240			720
	GTA ATG CTT GAG AGT CCA AGT GAG GTT AAT AGA GCA ATG GAC GAA TTC Val Met Leu Glu Ser Pro Ser Glu Val Asn Arg Ala Met Asp Glu Phe 245 250 255			768
	ATT TCT TCA GCA CAG TTC TAA			789

Ile Ser Ser Ala Gln Phe
260

5

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 756 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

20

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...753

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

25

	TTG AGA TTG AGG AAA TTT GAA GAG ATA AAC CTC GTT CTT TCG GGA GGA	48
	Leu Arg Leu Arg Lys Phe Glu Glu Ile Asn Leu Val Leu Ser Gly Gly	
	1 5 10 15	
30	GCT GCA AAG GGC ATA GCC CAC ATA GGT GTT TTG AAA GCT ATA AAC GAG	96
	Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys Ala Ile Asn Glu	
	20 25 30	
35	CTC GGT ATA AGG GTG AGG GCT TTA AGC GGG GTG AGC GCC GGG GCA ATC	144
	Leu Gly Ile Arg Val Arg Ala Leu Ser Gly Val Ser Ala Gly Ala Ile	
	35 40 45	
40	GTT TCG GTC TTT TAT GCC TCA GGC TAC TCC CCT GAA GGG ATG TTC AGC	192
	Val Ser Val Phe Tyr Ala Ser Gly Tyr Ser Pro Glu Gly Met Phe Ser	
	50 55 60	
45	CTT CTG AAG AGG GTA AAC TGG CTG AAG CTG TTT AAG TTC AAG CCA CCT	240
	Leu Leu Lys Arg Val Asn Trp Leu Lys Leu Phe Lys Phe Lys Pro Pro	
	65 70 75 80	
50	CTG AAG GGA TTG ATA GGG TGG GAG AAG GCT ATA AGA TTC CTT GAG GAA	288
	Leu Lys Gly Leu Ile Gly Trp Glu Lys Ala Ile Arg Phe Leu Glu Glu	
	85 90 95	
55	GTT CTC CCT TAC AGG AGA ATA GAA AAA CTT GAG ATA CCG ACG TAT ATA	336
	Val Leu Pro Tyr Arg Arg Ile Glu Lys Leu Glu Ile Pro Thr Tyr Ile	
	100 105 110	
	TGC GCG ACG GAT TTA TAC TCG GGA AGG GCT CTA TAC CTC TCG GAA GGG	384
	Cys Ala Thr Asp Leu Tyr Ser Gly Arg Ala Leu Tyr Leu Ser Glu Gly	
	115 120 125	
55	AGT TTA ATC CCC GCA CTT CTC GGC AGC TGT GCA ATT CCC GGC ATA TTT	432
	Ser Leu Ile Pro Ala Leu Leu Gly Ser Cys Ala Ile Pro Gly Ile Phe	
	130 135 140	

EP 0 880 590 B1

5	GAA CCC GTT GAG TAT AAG AAT TAC TTG CTC GTT GAC GGA GGT ATA GTT Glu Pro Val Glu Tyr Lys Asn Tyr Leu Leu Val Asp Gly Gly Ile Val 145 150 155 160	480
10	AAC AAC CTT CCC GTT GAG CCC TTT CAG GAA AGC GGT ATT CCC ACC GTT Asn Asn Leu Pro Val Glu Pro Phe Gln Glu Ser Gly Ile Pro Thr Val 165 170 175	528
15	TGC GTT GAT GTC CTT CCC ATA GAG CCG GAA AAG GAT ATA AAG AAC ATT Cys Val Asp Val Leu Pro Ile Glu Pro Glu Lys Asp Ile Lys Asn Ile 180 185 190	576
20	CTT CAC ATC CTT TTG AGG AGC TTC TTT CTT GCG GTC CGC TCA AAC TCC Leu His Ile Leu Leu Arg Ser Phe Phe Leu Ala Val Arg Ser Asn Ser 195 200 205	624
25	GAA AAG AGA AAG GAG TTT TGT GAC CTC GTT ATA GTT CCT GAG CTT GAG Glu Lys Arg Lys Glu Phe Cys Asp Leu Val Ile Val Pro Glu Leu Glu 210 215 220	672
30	GAG TTC ACA CCC CTT GAT GTT AGA AAA GCG GAC CAA ATA ATG GAG AGG Glu Phe Thr Pro Leu Asp Val Arg Lys Ala Asp Gln Ile Met Glu Arg 225 230 235 240	720
35	GGA TAC ATA AAG GCC TTA GAG GTA CTT TCT GAA TAG Gly Tyr Ile Lys Ala Leu Glu Val Leu Ser Glu 245 250	756

30 (2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 894 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...891

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

50	ATG TTT AAT ATC AAT GTC TTT GTT AAT ATA TCT TGG CTG TAT TTT TCA Met Phe Asn Ile Asn Val Phe Val Asn Ile Ser Trp Leu Tyr Phe Ser 1 5 10 15	48
55	GGG ATA GTT ATG AAG ACT GTG GAA GAG TAT GCG CTA CTT GAA ACA GGC Gly Ile Val Met Lys Thr Val Glu Glu Tyr Ala Leu Leu Glu Thr Gly 20 25 30	96

EP 0 880 590 B1

5	GTA AGA GTG TTT TAT CGG TGT GTA ATC CCG GAG AAA GCT TTT AAC ACT Val Arg Val Phe Tyr Arg Cys Val Ile Pro Glu Lys Ala Phe Asn Thr 35 40 45	144
10	TTG ATA ATA GGT TCA CAC GGA TTG GGG GCG CAC AGT GGA ATC TAC ATT Leu Ile Ile Gly Ser His Gly Leu Gly Ala His Ser Gly Ile Tyr Ile 50 55 60	192
15	AGT GTT GCT GAA GAA TTT GCT AGG CAC GGA TTT GGA TTC TGC ATG CAC Ser Val Ala Glu Glu Phe Ala Arg His Gly Phe Gly Phe Cys Met His 65 70 75 80	240
20	GAT CAA AGG GGA CAT GGG AGA ACG GCA AGC GAT AGA GAA AGA GGG TAT Asp Gln Arg Gly His Gly Arg Thr Ala Ser Asp Arg Glu Arg Gly Tyr 85 90 95	288
25	GTG GAG GGC TTT CAC AAC TTC ATA GAG GAT ATG AAG GCC TTC TCC GAT Val Glu Gly Phe His Asn Phe Ile Glu Asp Met Lys Ala Phe Ser Asp 100 105 110	336
30	TAT GCC AAG TGG CGC GTG GGA GGT GAC GAA ATA ATA TTG CTA GGA CAC Tyr Ala Lys Trp Arg Val Gly Gly Asp Glu Ile Ile Leu Leu Gly His 115 120 125	384
35	AGT ATG GGC GGG CTG ATA GCG CTC TTA ACA GTT GCA ACT TAT AAA GAA Ser Met Gly Gly Leu Ile Ala Leu Leu Thr Val Ala Thr Tyr Lys Glu 130 135 140	432
40	ATC GCC AAG GGA GTT ATC GCG CTA GCC CCG GCC CTC CAA ATC CCC TTA Ile Ala Lys Gly Val Ile Ala Leu Ala Pro Ala Leu Gln Ile Pro Leu 145 150 155 160	480
45	ACC CCG GCT AGA AGA CTT GTT CTA AGC CTC GCG TCA AGG CTT GCC CCG Thr Pro Ala Arg Arg Leu Val Leu Ser Leu Ala Ser Arg Leu Ala Pro 165 170 175	528
50	CAT TCT AAG ATC ACC TTA CAA AGG AGA TTG CCG CAG AAA CCA GAG GGT His Ser Lys Ile Thr Leu Gln Arg Arg Leu Pro Gln Lys Pro Glu Gly 180 185 190	576
55	TTT CAA AGA GCA AAA GAT ATA GAA TAC AGT CTG AGT GAA ATA TCA GTC Phe Gln Arg Ala Lys Asp Ile Glu Tyr Ser Leu Ser Glu Ile Ser Val 195 200 205	624
60	AAG CTC GTG GAC GAA ATG ATT AAA GCA TCA TCT ATG TTC TGG ACC ATA Lys Leu Val Asp Glu Met Ile Lys Ala Ser Ser Met Phe Trp Thr Ile 210 215 220	672
65	GCA GGG GAA ATT AAT ACT CCC GTC CTG CTT ATT CAT GGG GAA AAA GAC Ala Gly Glu Ile Asn Thr Pro Val Leu Leu Ile His Gly Glu Lys Asp 225 230 235 240	720
70	AAT GTC ATA CCT CCG GAG GCG AGC AAA AAA GCC TAC CAA TTA ATA CCT Asn Val Ile Pro Pro Glu Ala Ser Lys Ala Tyr Gln Leu Ile Pro	768

EP 0 880 590 B1

	245	250	255	
5	TCA TTC CCT AAA GAG TTG AAA ATA TAC CCC GAT CTT GGA CAC AAC TTG Ser Phe Pro Lys Glu Leu Lys Ile Tyr Pro Asp Leu Gly His Asn Leu 260	265	270	816
10	TTT TTT GAA CCA GGC GCG GTG AAA ATC GTC ACA GAC ATT GTA GAG TGG Phe Phe Glu Pro Gly Ala Val Lys Ile Val Thr Asp Ile Val Glu Trp 275	280	285	864
15	GTT AAG AAT CTA CCC AGG GAA AAT CCT TAA Val Lys Asn Leu Pro Arg Glu Asn Pro 290	295		894
20	(2) INFORMATION FOR SEQ ID NO:28:			
	(i) SEQUENCE CHARACTERISTICS:			
25	(A) LENGTH: 789 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear			
30	(ii) MOLECULE TYPE: Genomic DNA (ix) FEATURE:			
	(A) NAME/KEY: Coding Sequence (B) LOCATION: 1...786			
35	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:			
	ATG GAG GTT TAC AAG GCC AAA TTC GGC GAA GCA AAG CTC GGC TGG GTC Met Glu Val Tyr Lys Ala Lys Phe Gly Glu Ala Lys Leu Gly Trp Val 1 5 10 15			48
40	GTT CTG GTT CAT GGC CTC GGC GAG CAC AGC GGA AGG TAT GGA AGA CTG Val Leu Val His Gly Leu Gly Glu His Ser Gly Arg Tyr Gly Arg Leu 20 25 30			96
45	ATT AAG GAA CTC AAC TAT GCC GGC TTT GGA GTT TAC ACC TTC GAC TGG Ile Lys Glu Leu Asn Tyr Ala Gly Phe Gly Val Tyr Thr Phe Asp Trp 35 40 45			144
50	CCC GGC CAC GGG AAG AGC CCG GGC AAG AGA GGG CAC ACG AGC GTC GAG Pro Gly His Gly Lys Ser Pro Gly Lys Arg Gly His Thr Ser Val Glu 50 55 60			192
55	GAG GCG ATG GAA ATC ATC GAC TCG ATA ATC GAG GAG ATC AGG GAG AAG Glu Ala Met Glu Ile Ile Asp Ser Ile Ile Glu Glu Ile Arg Glu Lys 65 70 75 80			240
	CCC TTC CTC TTC GGC CAC AGC CTC GGT GGT CTA ACT GTC ATC AGG TAC Pro Phe Leu Phe Gly His Ser Leu Gly Gly Leu Thr Val Ile Arg Tyr 85 90 95			288
55	GCT GAG ACG CGG CCC GAT AAA ATA CGG GGA TTA ATA GCT TCC TCG CCT			336

EP 0 880 590 B1

	Ala Glu Thr Arg Pro Asp Lys Ile Arg Gly Leu Ile Ala Ser Ser Pro			
	100	105	110	
5	GCC CTC GCC AAG AGC CCG GAA ACG CCG GGC TTC ATG GTG GCC CTC GCG Ala Leu Ala Lys Ser Pro Glu Thr Pro Gly Phe Met Val Ala Leu Ala		384	
	115	120	125	
10	AAG TTC CTT GGA AAG ATC GCC CCG GGA GTT GTT CTC TCC AAC GGC ATA Lys Phe Leu Gly Lys Ile Ala Pro Gly Val Val Leu Ser Asn Gly Ile		432	
	130	135	140	
15	AAG CCG GAA CTC CTC TCG AGG AAC AGG GAC GCC GTG AGG AGG TAC GTT Lys Pro Glu Leu Leu Ser Arg Asn Arg Asp Ala Val Arg Arg Tyr Val		480	
	145	150	155	160
	GAA GAC CCA CTC GTC CAC GAC AGG ATT TCG GCC AAG CTG GGA AGG AGC Glu Asp Pro Leu Val His Asp Arg Ile Ser Ala Lys Leu Gly Arg Ser		528	
	165	170	175	
20	ATC TTC GTG AAC ATG GAG CTG GCC CAC AGG GAG GCG GAC AAG ATA AAA Ile Phe Val Asn Met Glu Leu Ala His Arg Glu Ala Asp Lys Ile Lys		576	
	180	185		
25	180	190		
	GTC CCG ATC CTC CTT CTG ATC GGC ACT GGC GAT GTA ATA ACC CCG CCT Val Pro Ile Leu Leu Leu Ile Gly Thr Gly Asp Val Ile Thr Pro Pro		624	
	195	200	205	
30	GAA GGC TCA CGC AGA CTC TTC GAG GAG CTG GCC GTC GAG AAC AAA ACC Glu Gly Ser Arg Arg Leu Phe Glu Glu Leu Ala Val Glu Asn Lys Thr		672	
	210	215	220	
	210	215	220	
35	CTG AGG GAG TTC GAG GGG GCG TAC CAC GAG ATA TTT GAA GAC CCC GAG Leu Arg Glu Phe Glu Gly Ala Tyr His Glu Ile Phe Glu Asp Pro Glu		720	
	225	230	235	240
	225	230	235	240
	TGG GCC GAG GAG TTC CAC GAA ACA ATT GTT AAG TGG CTG GTT GAA AAA Trp Ala Glu Glu Phe His Glu Thr Ile Val Lys Trp Leu Val Glu Lys		768	
	245	250	255	
40	TCG TAC TCT TCG GCT CAA TAA Ser Tyr Ser Ser Ala Gln		789	
	260			

(2) INFORMATION FOR SEQ ID NO:29:

45 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 750 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- 50 (D) TOPOLOGY: linear

55 (ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...747

EP 0 880 590 B1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

5	TTG ATT GGC AAT TTG AAA TTG AAG AGG TTT GAA GAG GTT AAC TTA GTT Leu Ile Gly Asn Leu Lys Leu Lys Arg Phe Glu Glu Val Asn Leu Val 1 5 10 15	48
10	CTT TCG GGA GGG GCT GCC AAG GGT ATC GCC CAT ATA GGT GTT TTA AAA Leu Ser Gly Gly Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys 20 25 30	96
15	GCT CTG GAA GAG CTC GGT ATA AAG GTA AAG AGG CTC AGC GGG GTA AGT Ala Leu Glu Leu Gly Ile Lys Val Lys Arg Leu Ser Gly Val Ser 35 40 45	144
20	GCT GGA GCT ATC GTT TCC GTC TTT TAC GCT TCG GGC TAC ACT CCC GAC Ala Gly Ala Ile Val Ser Val Phe Tyr Ala Ser Gly Tyr Thr Pro Asp 50 55 60	192
25	GAG ATG TTA AAA CTC CTG AAA GAG GTA AAC TGG CTC AAA CTT TTT AAG Glu Met Leu Lys Leu Lys Glu Val Asn Trp Leu Lys Leu Phe Lys 65 70 75 80	240
30	TTC AAA ACA CCG AAA ATG GGC TTA ATG GGG TGG GAG AAG GCT GCA GAG Phe Lys Thr Pro Lys Met Gly Leu Met Gly Trp Glu Lys Ala Ala Glu 85 90 95	288
35	TTT TTG GAA AAA GAG CTC GGA GTT AAG AGG CTG GAA GAC CTG AAC ATA Phe Leu Glu Lys Glu Leu Gly Val Lys Arg Leu Glu Asp Leu Asn Ile 100 105 110	336
40	CCA ACC TAT CTT TGC TCG GCG GAT CTG TAC ACG GGA AAG GCT CTT TAC Pro Thr Tyr Leu Cys Ser Ala Asp Leu Tyr Thr Gly Lys Ala Leu Tyr 115 120 125	384
45	TTC GGC AGA GGT GAC TTA ATT CCC GTG CTT CTC GGA AGT TGT TCC ATA Phe Gly Arg Gly Asp Leu Ile Pro Val Leu Leu Gly Ser Cys Ser Ile 130 135 140	432
50	CCC GGG ATT TTT GAA CCA GTT GAG TAC GAG AAT TTT CTA CTT GTT GAC Pro Gly Ile Phe Glu Pro Val Glu Tyr Glu Asn Phe Leu Leu Val Asp 145 150 155 160	480
55	GGA GGT ATA GTG AAC AAC CTG CCC GTA GAA CCT TTG GAA AAG TTC AAA Gly Gly Ile Val Asn Asn Leu Pro Val Glu Pro Leu Glu Lys Phe Lys 165 170 175	528
60	GAA CCC ATA ATC GGG GTA GAT GTG CTT CCC ATA ACT CAA GAA AGA AAG Glu Pro Ile Ile Gly Val Asp Val Leu Pro Ile Thr Gln Glu Arg Lys 180 185 190	576
65	ATT AAA AAT ATA CTC CAC ATC CTT ATA AGG AGC TTC TTT CTG GCG GTT Ile Lys Asn Ile Leu His Ile Leu Ile Arg Ser Phe Phe Leu Ala Val 195 200 205	624

50

55

CGT TCC AAT TCG GAA AAG AGA AAG GAG TTC TGC AAC GTA GTT ATA GAA	672
Arg Ser Asn Ser Glu Lys Arg Lys Glu Phe Cys Asn Val Val Ile Glu	
210 215 220	
CCT CCC CTT GAA GAG TTC TCT CCT CTG GAC GTA AAT AAG GCG GAC GAG	720
Pro Pro Leu Glu Glu Phe Ser Pro Leu Asp Val Asn Lys Ala Asp Glu	
225 230 235 240	
ATA TTC TGC GGG GAT ATG AGA GCA CTT TAA	750
Ile Phe Cys Gly Asp Met Arg Ala Leu	
245	

15 (2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1017 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...1014

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

35	ATG CCA GCT AAT GAC TCA CCC ACG ATC GAC TTT AAT CCT CGC GGC ATT Met Pro Ala Asn Asp Ser Pro Thr Ile Asp Phe Asn Pro Arg Gly Ile 1 5 10 15	48
40	CTT CGC AAC GCT CAC GCA CAG GTT ATT TTA GCG ACT TCC GGC TTG CGC Leu Arg Asn Ala His Ala Gln Val Ile Leu Ala Thr Ser Gly Leu Arg 20 25 30	96
45	AAA GCG TTT TTG AAA CGC ACG CAC AAG AGC TAC CTC AGC ACT GCC CAA Lys Ala Phe Leu Lys Arg Thr His Lys Ser Tyr Leu Ser Thr Ala Gln 35 40 45	144
50	TGG CTG GAG CTC GAT GCC GGC AAC GGA GTT ACC TTG GCC GGA GAG CTT Trp Leu Glu Leu Asp Ala Gly Asn Gly Val Thr Leu Ala Gly Glu Leu 50 55 60	192
55	AAC ACA GCG CCT GCA ACT GCA TCC TCC CAC CCG GCG CAC AAG AAC Asn Thr Ala Pro Ala Thr Ala Ser Ser Ser His Pro Ala His Lys Asn 65 70 75 80	240
55	ACT CTG GTT ATT GTG CTG CAC GGC TGG GAA GGC TCC AGC CAG TCG GCC Thr Leu Val Ile Val Leu His Gly Trp Glu Gly Ser Ser Gln Ser Ala 85 90 95	288

EP 0 880 590 B1

	TAT GCG ACC TCC GCT GGC AGC ACG CTT TTC GAC AAT GGG TTC GAC ACT Tyr Ala Thr Ser Ala Gly Ser Thr Leu Phe Asp Asn Gly Phe Asp Thr 100 105 110	336
5	TTT CGC CTT AAT TTT CGC GAT CAC GGC GAC ACC TAC CAC TTA AAC CGC Phe Arg Leu Asn Phe Arg Asp His Gly Asp Thr Tyr His Leu Asn Arg 115 120 125	384
10	GCG ATA TTT AAC TCA TCG CTG ATT GAC GAA GTA GTG GGC GCA GTC AAA Gly Ile Phe Asn Ser Ser Leu Ile Asp Glu Val Val Gly Ala Val Lys 130 135 140	432
15	GCC ATC CAG CAG CAA ACC GAC TAC GAC AAG TAT TGC CTG ATG GGG TTC Ala Ile Gln Gln Gln Thr Asp Tyr Asp Lys Tyr Cys Leu Met Gly Phe 145 150 155 160	480
	TCA CTG GGT GGG AAC TTT GCC TTG CGC GTC GCG GTG CGG GAA CAG CAT Ser Leu Gly Gly Asn Phe Ala Leu Arg Val Ala Val Arg Glu Gln His 165 170 175	528
20	CTC GCT AAA CCG CTA GCG GGC GTG CTC GCC GTA TGC CCG GTA CTC GAC Leu Ala Lys Pro Leu Ala Gly Val Leu Ala Val Cys Pro Val Leu Asp 180 185 190	576
25	CCC GCA CAC ACC ATG ATG GCC CTA AAC CGA GGT GCG TTT TTC TAC GGC Pro Ala His Thr Met Met Ala Leu Asn Arg Gly Ala Phe Phe Tyr Gly 195 200 205	624
30	CGC TAT TTT GCG CAT AAA TGG AAG CGC TCG TTA ACC GCA AAA CTT GCA Arg Tyr Phe Ala His Lys Trp Lys Arg Ser Leu Thr Ala Lys Leu Ala 210 215 220	672
	GCT TTC CCA GAC TAC AAA TAC GGC AAA GAT TTA AAA TCG ATA CAC ACG Ala Phe Pro Asp Tyr Lys Tyr Gly Lys Asp Leu Lys Ser Ile His Thr 225 230 235 240	720
35	CTT GAT GAG TTA AAC AAC TAT TTC ATT CCC CGC TAC ACC GGC TTC AAC Leu Asp Glu Leu Asn Asn Tyr Phe Ile Pro Arg Tyr Thr Gly Phe Asn 245 250 255	768
40	TCA GTC TCC GAA TAC TTC AAA AGT TAC ACG CTC ACC GGG CAG AAG CTC Ser Val Ser Glu Tyr Phe Lys Ser Tyr Thr Leu Thr Gly Gln Lys Leu 260 265 270	816
	GGC TTT CTC AAC TGC CCC AGT TAC ATT CTG GCA GCT GGC GAC GAC CCA Ala Phe Leu Asn Cys Pro Ser Tyr Ile Leu Ala Ala Gly Asp Asp Pro 275 280 285	864
45	ATA ATT CCA GCA TCC GAC TTT CAG AAA ATA GCC AAG CCT GCG AAT CTG Ile Ile Pro Ala Ser Asp Phe Gln Lys Ile Ala Lys Pro Ala Asn Leu 290 295 300	912
50	CAC ATA ACA GTA ACG CAA CAA GGT TCT CAT TGC GCA TAC CTG GAA AAC His Ile Thr Val Thr Gln Gln Gly Ser His Cys Ala Tyr Leu Glu Asn 305 310 315 320	960

5 CTG CAT AAA CCT AGT GCT GCC GAC AAA TAT GCG GTG AAA TTA TTT GGA 1008
 Leu His Lys Pro Ser Ala Ala Asp Lys Tyr Ala Val Lys Leu Phe Gly
 325 330 335

GCC TGT TGA 1017
 Ala Cys

10

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 936 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

20

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- 25 (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 1...933

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

30

ATG CTT GAT ATG CCA ATC GAC CCT GTT TAC TAC CAG CTT GCT GAG TAT 48
 Met Leu Asp Met Pro Ile Asp Pro Val Tyr Tyr Gln Leu Ala Glu Tyr
 1 5 10 15

35

TTC GAC AGT CTG CCG AAG TTC GAC CAG TTT TCC TCG GCC AGA GAG TAC 96
 Phe Asp Ser Leu Pro Lys Phe Asp Gln Phe Ser Ser Ala Arg Glu Tyr
 20 25 30
 AGG GAG GCG ATA AAT CGA ATA TAC GAG GAG AGA AAC CGG CAG CTG AGC 144
 Arg Glu Ala Ile Asn Arg Ile Tyr Glu Glu Arg Asn Arg Gln Leu Ser
 35 40 45

40

CAG CAT GAG AGG GTT GAA AGA GTT GAG GAC AGG ACG ATT AAG GGG AGG 192
 Gln His Glu Arg Val Glu Arg Val Glu Asp Arg Thr Ile Lys Gly Arg
 50 55 60

45

AAC GGA GAC ATC AGA GTC AGA GTT TAC CAG CAG AAG CCC GAT TCC CCG 240
 Asn Gly Asp Ile Arg Val Arg Val Tyr Gln Gln Lys Pro Asp Ser Pro
 65 70 75 80

50

GGT CTG GTT TAC TAT CAC GGT GGT GGA TTT GTG ATT TGC AGC ATC GAG 288
 Val Leu Val Tyr Tyr His Gly Gly Gly Phe Val Ile Cys Ser Ile Glu
 85 90 95

55

TCG CAC GAC GCC TTA TGC AGG AGA AYY GCG AGA CTT TCA AAC TCT ACC 336
 Ser His Asp Ala Leu Cys Arg Arg Ile Ala Arg Leu Ser Asn Ser Thr
 100 105 110

GTA GTC TCC GTG GAT TAC AGG CTC GCT CCT GAG CAC AAG TTT CCC CCC 384
 Val Val Ser Val Asp Tyr Arg Leu Ala Pro Glu His Phe Pro Ala

EP 0 880 590 B1

	115	120	125	
5	CCA GTT TAT CAT TGC TAC GAT GCG ACC AAG TGG GTT GCT GAG AAC CGG Ala Val Tyr Asp Cys Tyr Asp Ala Thr Lys Trp Val Ala Glu Asn Ala 130 135 140			432
10	GAG GAG CTG AGG ATT GAC CCG TCA AAA ATC TTC GTT GGG GGG GAC AGT Glu Glu Leu Arg Ile Asp Pro Ser Lys Ile Phe Val Gly Gly Asp Ser 145 150 155 160			480
15	GCG GGA CG AAT CTT GCC CGG GCG CTT TCA ATA ATG GCG AGA GAC AGC Ala Gly Gly Asn Leu Ala Ala Val Ser Ile Met Ala Arg Asp Ser 165 170 175			528
20	GGA GAA GAT TTC ATA AAG CAT CAA ATT CTA ACT TAC CCC GTT GTG AAC Gly Glu Asp Phe Ile Lys His Gln Ile Leu Ile Tyr Pro Val Val Asn 180 185 190			576
25	TTT GTA GCC CCC ACA CCA TCG CTT CTG GAG TTT GGA GAG GGG CTG TGG Phe Val Ala Pro Thr Pro Ser Leu Leu Glu Phe Gly Glu Gly Leu Trp 195 200 205			624
30	ATT CTC GAC CAG AAG ATA ATG AGT TGG TTC TCG GAG CAG TAC TTC TCC Ile Leu Asp Gln Lys Ile Met Ser Trp Phe Ser Glu Gln Tyr Phe Ser 210 215 220			672
35	AGA GAG GAA GAT AAG TTC AAG CCC CTC GCC TCC GTA ATC TTT GCG GAC Arg Glu Glu Asp Lys Phe Asn Pro Leu Ala Ser Val Ile Phe Ala Asp 225 230 235 240			720
40	CTT GAG AAC CTA CCT CCT GCG CTG ATC ATA ACC GCC GAA TAC GAC CCG Leu Glu Asn Leu Pro Pro Ala Leu Ile Ile Thr Ala Glu Tyr Asp Pro 245 250 255			768
45	CTG AGA GAT GAA GGA GAA GTT TTC GGG CAG ATG CTG AGA AGA GAC GCC GGT Leu Arg Asp Glu Gly Glu Val Phe Gly Gln Met Leu Arg Arg Ala Gly 260 265 270			816
50	GTT GAG GCG AGC ATC GTC AGA TAC AGA GGC GTG CTT CAC GGA TTC ATC Val Glu Ala Ser Ile Val Arg Tyr Arg Gly Val Leu His Gly Phe Ile 275 280 285			864
55	AAT TAC TAT CCC GTG CTG AAG GCT GCG AGG GAT GCG ATA AAC CAG ATT Asn Tyr Tyr Pro Val Leu Lys Ala Ala Arg Asp Ala Ile Asn Gln Ile 290 295 300			912
	GCC GCT CTT CTT GTG TTC GAC TAG Ala Ala Leu Leu Val Phe Asp 305 310			936

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 918 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

EP 0 880 590 B1

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
 (B) LOCATION: 1...915

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

10	ATG CCC CTA GAT CCT AGA ATT AAA AAG TTA CTA GAA TCA GCT CTT ACT Met Pro Leu Asp Pro Arg Ile Lys Lys Leu Leu Glu Ser Ala Leu Thr 1 5 10 15	48
15	ATA CCA ATT GGT AAA GCC CCA GTA GAA GAG GTA AGA AAG ATA TTT AGG Ile Pro Ile Gly Lys Ala Pro Val Glu Glu Val Arg Lys Ile Phe Arg 20 25 30	96
20	CAA TTA GCG TCG GCA GCT CCC AAA GTC GAA GTT GGA AAA GTA GAA GAT Gln Leu Ala Ser Ala Ala Pro Lys Val Glu Val Gly Lys Val Glu Asp 35 40 45	144
25	ATA AAA ATA CCA GGC AGT GAA ACC GTT ATA AAC GCT AGA GTG TAT TTT Ile Lys Ile Pro Gly Ser Glu Thr Val Ile Asn Ala Arg Val Tyr Phe 50 55 60	192
30	CCG AAG AGT AGC GGT CCT TAT GGT GTT CTA GTG TAT CTT CAT GGA GGC Pro Lys Ser Ser Gly Pro Tyr Gly Val Leu Val Tyr Leu His Gly Gly 65 70 75 80	240
35	GGT TTT GTA ATA GGC GAT GTG GAA TCT TAT GAC CCA TTA TGT AGA GCA Gly Phe Val Ile Gly Asp Val Glu Ser Tyr Asp Pro Leu Cys Arg Ala 85 90 95	288
40	ATT ACA AAT GCG TGC AAT TGC GTT GTA GTA TCA GTG GAC TAT AGG TTA Ile Thr Asn Ala Cys Asn Cys Val Val Ser Val Asp Tyr Arg Leu 100 105 110	336
45	GCT CCA GAA TAC AAG TTT CCT TCT GCA GTT ATC GAT TCA TTT GAC GCT Ala Pro Glu Tyr Lys Phe Pro Ser Ala Val Ile Asp Ser Phe Asp Ala 115 120 125	384
50	ACT AAT TGG GTT TAT AAC AAT TTA GAT AAA TTT GAT GGA AAG ATG GGA Thr Asn Trp Val Tyr Asn Asn Leu Asp Lys Phe Asp Gly Lys Met Gly 130 135 140	432
55	GTT GCG ATT GCG GGA GAT AGT GCT GGA GGA AAT TTG GCA GCG GTT GTA Val Ala Ile Ala Gly Asp Ser Ala Gly Gly Asn Leu Ala Ala Val Val 145 150 155 160	480
60	GCT CTT CTT TCA AAG GGT AAA ATT AAT TTG AAG TAT CAA ATA CTG GTT Ala Leu Leu Ser Lys Gly Lys Ile Asn Leu Lys Tyr Gln Ile Leu Val 165 170 175	528

50

55

EP 0 880 590 B1

	TAC CCA GCG GTA AGT TTA GAT AAC GTT TCA AGA TCC ATG ATA GAG TAC Tyr Pro Ala Val Ser Leu Asp Asn Val Ser Arg Ser Met Ile Glu Tyr 180 185 190	576
5	TCT GAT GGG TTC TTC CTT ACC AGA GAG CAT ATA GAG TGG TTC GGT TCT Ser Asp Gly Phe Phe Leu Thr Arg Glu His Ile Glu Trp Phe Gly Ser 195 200 205	624
10	CAA TAC TTA CGA AGC CCT GCA GAT TTG CTA GAC TTT AGG TTC TCT CCA Gln Tyr Leu Arg Ser Pro Ala Asp Leu Leu Asp Phe Arg Phe Ser Pro 210 215 220	672
15	ATT CTG GCG CAA GAT TTC AAC GGA TTA CCT CCA GCC TTG ATA ATA ACA Ile Leu Ala Gln Asp Phe Asn Gly Leu Pro Pro Ala Leu Ile Ile Thr 225 230 235 240	720
	GCA GAA TAC GAT CCA CTA AGG GAT CAA GGA GAA GCG TAT GCA AAT AAA Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Ala Tyr Ala Asn Lys 245 250 255	768
20	CTA CTA CAA GCT GGA GTC TCA GTT ACT AGT GTG AGA TTT AAC AAC GTT Leu Leu Gln Ala Gly Val Ser Val Thr Ser Val Arg Phe Asn Asn Val 260 265 270	816
25	ATA CAC GGA TTC CTC TCA TTC TTT CCG TTG ATG GAG CAA GGA AGA GAT Ile His Gly Phe Leu Ser Phe Pro Leu Met Glu Gln Gly Arg Asp 275 280 285	864
	GCT ATA GGT CTG ATA GGG TCT GTG TTA AGA CGA GTA TTT TAT GAT AAA Ala Ile Gly Leu Ile Gly Ser Val Leu Arg Arg Val Phe Tyr Asp Lys 290 295 300	912
30	ATT TAA Ile 305	918

35 (2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 184 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

45 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

50	Met Ser Leu Asn Lys His Ser Trp Met Asp Met Ile Ile Phe Ile Leu 1 5 10 15
	Ser Phe Ser Phe Pro Leu Thr Met Ile Ala Leu Ala Ile Ser Met Ser 20 25 30
55	Ser Trp Phe Asn Ile Trp Asn Asn Ala Leu Ser Asp Leu Gly His Ala 35 40 45

Val Lys Ser Ser Val Ala Pro Ile Phe Asn Leu Gly Leu Ala Ile Gly
 50 55 60
 Gly Ile Leu Ile Val Ile Val Gly Leu Arg Asn Leu Tyr Ser Trp Ser
 65 70 75 80
 Arg Val Lys Gly Ser Leu Ile Ile Ser Met Gly Val Phe Leu Asn Leu
 85 90 95
 Ile Gly Val Phe Asp Glu Val Tyr Gly Trp Ile His Phe Leu Val Ser
 100 105 110
 Val Leu Phe Phe Leu Ser Ile Ile Ala Tyr Phe Ile Ala Ile Ser Ile
 115 120 125
 Leu Asp Lys Ser Trp Ile Ala Val Leu Leu Ile Ile Gly His Ile Ala
 130 135 140
 Met Trp Tyr Leu His Phe Ala Ser Glu Ile Pro Arg Gly Ala Ala Ile
 145 150 155 160
 Pro Glu Leu Leu Ala Val Phe Ser Phe Leu Pro Phe Tyr Ile Arg Asp
 165 170 175
 Tyr Phe Lys Ser Tyr Thr Lys Arg
 180

20

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 346 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 30 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

35 Met Lys Leu Leu Glu Pro Thr Asn Thr Ser Tyr Thr Leu Leu Gln Asp
 1 5 10 15
 Leu Ala Leu His Phe Ala Phe Tyr Trp Phe Leu Ala Val Tyr Thr Trp
 20 25 30
 40 Leu Pro Gly Val Leu Val Arg Gly Val Ala Val Asp Thr Gly Val Ala
 35 40 45
 Arg Val Pro Gly Leu Gly Arg Arg Gly Lys Arg Leu Leu Ala Ala
 50 55 60
 Val Ala Val Leu Ala Leu Val Val Ser Val Val Val Pro Ala Tyr Val
 65 70 75 80
 45 Ala Tyr Ser Ser Leu His Pro Glu Ser Cys Arg Pro Val Ala Pro Glu
 85 90 95
 Gly Leu Thr Tyr Lys Glu Phe Ser Val Thr Ala Glu Asp Gly Leu Val
 100 105 110
 50 Val Arg Gly Trp Val Leu Gly Pro Gly Ala Gly Gly Asn Pro Val Phe
 115 120 125
 55 Val Leu Met His Gly Tyr Thr Gly Cys Arg Ser Ala Pro Tyr Met Ala
 130 135 140
 60 Val Leu Ala Arg Glu Leu Val Glu Trp Gly Tyr Pro Val Val Val Phe
 145 150 155 160
 65 Asp Phe Arg Gly His Gly Glu Ser Gly Gly Ser Thr Thr Ile Gly Pro
 165 170 175

Arg Glu Val Leu Asp Ala Arg Ala Val Val Gly Tyr Val Ser Glu Arg
 180 185 190
 Phe Pro Gly Arg Arg Ile Ile Leu Val Gly Phe Ser Met Gly Gly Ala
 195 200 205
 5 Val Ala Ile Val Glu Gly Ala Gly Asp Pro Arg Val Tyr Ala Val Ala
 210 215 220
 Ala Asp Ser Pro Tyr Tyr Arg Leu Arg Asp Val Ile Pro Arg Trp Leu
 225 230 235 240
 10 Glu Tyr Lys Thr Pro Leu Pro Gly Trp Val Gly Val Leu Ala Gly Phe
 245 250 255
 Tyr Gly Arg Leu Met Ala Gly Val Asp Leu Gly Phe Gly Pro Ala Gly
 260 265 270
 Val Glu Arg Val Asp Lys Pro Leu Leu Val Val Tyr Gly Pro Arg Asp
 275 280 285
 15 Pro Leu Val Thr Arg Asp Glu Ala Arg Ser Leu Ala Ser Arg Ser Pro
 290 295 300
 Cys Gly Arg Leu Val Glu Val Pro Gly Ala Gly His Val Glu Ala Val
 305 310 315 320
 Asp Val Leu Gly Pro Gly Arg Tyr Ala Asp Met Leu Ile Glu Leu Ala
 325 330 335
 20 His Glu Glu Cys Pro Pro Gly Ala Gly Gly
 340 345

(2) INFORMATION FOR SEQ ID NO:35:

- 25
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 262 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear - (ii) MOLECULE TYPE: protein
 - (v) FRAGMENT TYPE: internal
- 35
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Met Pro Tyr Val Arg Asn Gly Gly Val Asn Ile Tyr Tyr Glu Leu Val
 40 1 5 10 15
 Asp Gly Pro Glu Pro Pro Ile Val Phe Val His Gly Trp Thr Ala Asn
 20 25 30
 Met Asn Phe Trp Lys Glu Gln Arg Arg Tyr Phe Ala Gly Arg Asn Met
 35 40 45
 Met Leu Phe Val Asp Asn Arg Gly His Gly Arg Ser Asp Lys Pro Leu
 45 50 55 60
 Gly Tyr Asp Phe Tyr Arg Phe Glu Asn Phe Ile Ser Asp Leu Asp Ala
 65 70 75 80
 Val Val Arg Glu Thr Gly Val Glu Lys Phe Val Leu Val Gly His Ser
 85 90 95
 50 Phe Gly Thr Met Ile Ser Met Lys Tyr Cys Ser Glu Tyr Arg Asn Arg
 100 105 110
 Val Leu Ala Leu Ile Leu Ile Gly Gly Ser Arg Ile Lys Leu Leu
 115 120 125
 His Arg Ile Gly Tyr Pro Leu Ala Lys Ile Leu Ala Ser Ile Ala Tyr
 130 135 140

Lys Lys Ser Ser Arg Leu Val Ala Asp Leu Ser Phe Gly Lys Asn Ala
 145 150 155 160
 Gly Glu Leu Lys Glu Trp Gly Trp Lys Gln Ala Met Asp Tyr Thr Pro
 165 170 175
 Ser Tyr Val Ala Met Tyr Thr Tyr Arg Thr Leu Thr Lys Val Asn Leu
 180 185 190
 Glu Asn Ile Leu Glu Lys Ile Asp Cys Pro Thr Leu Ile Ile Val Gly
 195 200 205
 Glu Glu Asp Ala Leu Leu Pro Val Ser Lys Ser Val Glu Leu Ser Arg
 210 215 220
 Arg Ile Glu Asn Ser Lys Leu Val Ile Ile Pro Asn Ser Gly His Cys
 225 230 235 240
 Val Met Leu Glu Ser Pro Ser Glu Val Asn Arg Ala Met Asp Glu Phe
 245 250 255
 Ile Ser Ser Ala Gln Phe
 260

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 251 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Leu Arg Leu Arg Lys Phe Glu Glu Ile Asn Leu Val Leu Ser Gly Gly
 1 5 10 15
 Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys Ala Ile Asn Glu
 20 25 30
 Leu Gly Ile Arg Val Arg Ala Leu Ser Gly Val Ser Ala Gly Ala Ile
 35 40 45
 Val Ser Val Phe Tyr Ala Ser Gly Tyr Ser Pro Glu Gly Met Phe Ser
 50 55 60
 Leu Leu Lys Arg Val Asn Trp Leu Lys Leu Phe Lys Phe Lys Pro Pro
 65 70 75 80
 Leu Lys Gly Leu Ile Gly Trp Glu Lys Ala Ile Arg Phe Leu Glu Glu
 85 90 95
 Val Leu Pro Tyr Arg Arg Ile Glu Lys Leu Glu Ile Pro Thr Tyr Ile
 100 105 110
 Cys Ala Thr Asp Leu Tyr Ser Gly Arg Ala Leu Tyr Leu Ser Glu Gly
 115 120 125
 Ser Leu Ile Pro Ala Leu Leu Gly Ser Cys Ala Ile Pro Gly Ile Phe
 130 135 140
 Glu Pro Val Glu Tyr Lys Asn Tyr Leu Leu Val Asp Gly Gly Ile Val
 145 150 155 160
 Asn Asn Leu Pro Val Glu Pro Phe Gln Glu Ser Gly Ile Pro Thr Val
 165 170 175
 Cys Val Asp Val Leu Pro Ile Glu Pro Glu Lys Asp Ile Lys Asn Ile
 180 185 190
 Leu His Ile Leu Leu Arg Ser Phe Phe Leu Ala Val Arg Ser Asn Ser
 195 200 205

Glu Lys Arg Lys Glu Phe Cys Asp Leu Val Ile Val Pro Glu Leu Glu
 210 215 220
 Glu Phe Thr Pro Leu Asp Val Arg Lys Ala Asp Gln Ile Met Glu Arg
 225 230 235 240
 Gly Tyr Ile Lys Ala Leu Glu Val Leu Ser Glu
 245 250

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 297 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Met Phe Asn Ile Asn Val Phe Val Asn Ile Ser Trp Leu Tyr Phe Ser
 1 5 10 15
 Gly Ile Val Met Lys Thr Val Glu Glu Tyr Ala Leu Leu Glu Thr Gly
 20 25 30
 Val Arg Val Phe Tyr Arg Cys Val Ile Pro Glu Lys Ala Phe Asn Thr
 35 40 45
 Leu Ile Ile Gly Ser His Gly Leu Gly Ala His Ser Gly Ile Tyr Ile
 50 55 60
 Ser Val Ala Glu Glu Phe Ala Arg His Gly Phe Gly Phe Cys Met His
 65 70 75 80
 Asp Gln Arg Gly His Gly Arg Thr Ala Ser Asp Arg Glu Arg Gly Tyr
 85 90 95
 Val Glu Gly Phe His Asn Phe Ile Glu Asp Met Lys Ala Phe Ser Asp
 100 105 110
 Tyr Ala Lys Trp Arg Val Gly Gly Asp Glu Ile Ile Leu Leu Gly His
 115 120 125
 Ser Met Gly Gly Leu Ile Ala Leu Leu Thr Val Ala Thr Tyr Lys Glu
 130 135 140
 Ile Ala Lys Gly Val Ile Ala Leu Ala Pro Ala Leu Gln Ile Pro Leu
 145 150 155 160
 Thr Pro Ala Arg Arg Leu Val Leu Ser Leu Ala Ser Arg Leu Ala Pro
 165 170 175
 His Ser Lys Ile Thr Leu Gln Arg Arg Leu Pro Gln Lys Pro Glu Gly
 180 185 190
 Phe Gln Arg Ala Lys Asp Ile Glu Tyr Ser Leu Ser Glu Ile Ser Val
 195 200 205
 Lys Leu Val Asp Glu Met Ile Lys Ala Ser Ser Met Phe Trp Thr Ile
 210 215 220
 Ala Gly Glu Ile Asn Thr Pro Val Leu Leu Ile His Gly Glu Lys Asp
 225 230 235 240
 Asn Val Ile Pro Pro Glu Ala Ser Lys Lys Ala Tyr Gln Leu Ile Pro
 245 250 255
 Ser Phe Pro Lys Glu Leu Lys Ile Tyr Pro Asp Leu Gly His Asn Leu
 260 265 270

Phe Phe Glu Pro Gly Ala Val Lys Ile Val Thr Asp Ile Val Glu Trp
 275 280 285
 Val Lys Asn Leu Pro Arg Glu Asn Pro
 290 295

5

(2) INFORMATION FOR SEQ ID NO:38:

10

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 262 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

15

- (ii) MOLECULE TYPE: protein

- (v) FRAGMENT TYPE: internal

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

	Met	Glu	Val	Tyr	Lys	Ala	Lys	Phe	Gly	Glu	Ala	Lys	Leu	Gly	Trp	Val	
25	1	5						10				15					
	Val	Leu	Val	His	Gly	Leu	Gly	Glu	His	Ser	Gly	Arg	Tyr	Gly	Arg	Leu	
				20				25				30					
	Ile	Lys	Glu	Leu	Asn	Tyr	Ala	Gly	Phe	Gly	Val	Tyr	Thr	Phe	Asp	Trp	
30		35					40				45						
	Pro	Gly	His	Gly	Lys	Ser	Pro	Gly	Lys	Arg	Gly	His	Thr	Ser	Val	Glu	
35		50			55				60								
	Glu	Ala	Met	Glu	Ile	Ile	Asp	Ser	Ile	Ile	Glu	Glu	Ile	Arg	Glu	Lys	
40		65			70				75			80					
	Pro	Phe	Leu	Phe	Gly	His	Ser	Leu	Gly	Gly	Leu	Thr	Val	Ile	Arg	Tyr	
45		85				90			95								
	Ala	Glu	Thr	Arg	Pro	Asp	Lys	Ile	Arg	Gly	Leu	Ile	Ala	Ser	Ser	Pro	
50		100			105				110								
	Ala	Leu	Ala	Lys	Ser	Pro	Glu	Thr	Pro	Gly	Phe	Met	Val	Ala	Leu	Ala	
55		115			120				125								
	Lys	Phe	Leu	Gly	Lys	Ile	Ala	Pro	Gly	Val	Val	Leu	Ser	Asn	Gly	Ile	
		130			135				140								
	Lys	Pro	Glu	Leu	Leu	Ser	Arg	Asn	Arg	Asp	Ala	Val	Arg	Arg	Tyr	Val	
60		145			150				155			160					
	Glu	Asp	Pro	Leu	Val	His	Asp	Arg	Ile	Ser	Ala	Lys	Leu	Gly	Arg	Ser	
65		165			170				175								
	Ile	Phe	Val	Asn	Met	Glu	Leu	Ala	His	Arg	Glu	Ala	Asp	Lys	Ile	Lys	
70		180			185				190								
	Val	Pro	Ile	Leu	Leu	Ile	Gly	Thr	Gly	Asp	Val	Ile	Thr	Pro	Pro		
75		195			200				205								
	Glu	Gly	Ser	Arg	Arg	Leu	Phe	Glu	Glu	Leu	Ala	Val	Glu	Asn	Lys	Thr	
80		210			215				220								
	Leu	Arg	Glu	Phe	Glu	Gly	Ala	Tyr	His	Glu	Ile	Phe	Glu	Asp	Pro	Glu	
85		225			230				235			240					
	Trp	Ala	Glu	Glu	Phe	His	Glu	Thr	Ile	Val	Lys	Trp	Leu	Val	Glu	Lys	
90		245			250				255								
	Ser	Tyr	Ser	Ser	Ala	Gln											
95		260															

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 249 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

10

	Leu	Ile	Gly	Asn	Leu	Lys	Leu	Lys	Arg	Phe	Glu	Glu	Val	Asn	Leu	Val	
	1				5				10					15			
	Leu	Ser	Gly	Gly	Ala	Ala	Lys	Gly	Ile	Ala	His	Ile	Gly	Val	Leu	Lys	
					20				25					30			
15		Ala	Leu	Glu	Glu	Leu	Gly	Ile	Lys	Val	Lys	Arg	Leu	Ser	Gly	Val	Ser
					35				40					45			
	Ala	Gly	Ala	Ile	Val	Ser	Val	Phe	Tyr	Ala	Ser	Gly	Tyr	Thr	Pro	Asp	
					50			55					60				
20	Glu	Met	Leu	Lys	Leu	Leu	Lys	Glu	Val	Asn	Trp	Leu	Lys	Leu	Phe	Lys	
					65			70			75			80			
	Phe	Lys	Thr	Pro	Lys	Met	Gly	Leu	Met	Gly	Trp	Glu	Lys	Ala	Ala	Glu	
					85				90					95			
	Phe	Leu	Glu	Lys	Glu	Leu	Gly	Val	Lys	Arg	Leu	Glu	Asp	Leu	Asn	Ile	
					100				105					110			
25	Pro	Thr	Tyr	Leu	Cys	Ser	Ala	Asp	Leu	Tyr	Thr	Gly	Lys	Ala	Leu	Tyr	
					115				120					125			
	Phe	Gly	Arg	Gly	Asp	Leu	Ile	Pro	Val	Leu	Gly	Ser	Cys	Ser	Ile		
					130				135					140			
30	Pro	Gly	Ile	Phe	Glu	Pro	Val	Glu	Tyr	Glu	Asn	Phe	Leu	Leu	Val	Asp	
					145				150			155			160		
	Gly	Gly	Ile	Val	Asn	Asn	Leu	Pro	Val	Glu	Pro	Leu	Glu	Phe	Lys		
					165				170					175			
	Glu	Pro	Ile	Ile	Gly	Val	Asp	Val	Leu	Pro	Ile	Thr	Gln	Glu	Arg	Lys	
					180				185					190			
35	Ile	Lys	Asn	Ile	Leu	His	Ile	Leu	Ile	Arg	Ser	Phe	Phe	Leu	Ala	Val	
					195				200					205			
	Arg	Ser	Asn	Ser	Glu	Lys	Arg	Lys	Glu	Phe	Cys	Asn	Val	Val	Ile	Glu	
					210				215					220			
	Pro	Pro	Leu	Glu	Glu	Phe	Ser	Pro	Leu	Asp	Val	Asn	Lys	Ala	Asp	Glu	
					225				230			235			240		
40	Ile	Phe	Cys	Gly	Asp	Met	Arg	Ala	Leu								
					245												

(2) INFORMATION FOR SEQ ID NO:40:

45

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 338 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Met Pro Ala Asn Asp Ser Pro Thr Ile Asp Phe Asn Pro Arg Gly Ile
 1 5 10 15
 Leu Arg Asn Ala His Ala Gln Val Ile Leu Ala Thr Ser Gly Leu Arg
 20 25 30
 Lys Ala Phe Leu Lys Arg Thr His Lys Ser Tyr Leu Ser Thr Ala Gln
 35 40 45
 Trp Leu Glu Leu Asp Ala Gly Asn Gly Val Thr Leu Ala Gly Glu Leu
 50 55 60
 Asn Thr Ala Pro Ala Thr Ala Ser Ser His Pro Ala His Lys Asn
 65 70 75 80
 Thr Leu Val Ile Val Leu His Gly Trp Glu Gly Ser Ser Gln Ser Ala
 85 90 95
 Tyr Ala Thr Ser Ala Gly Ser Thr Leu Phe Asp Asn Gly Phe Asp Thr
 100 105 110
 15 Phe Arg Leu Asn Phe Arg Asp His Gly Asp Thr Tyr His Leu Asn Arg
 115 120 125
 Gly Ile Phe Asn Ser Ser Leu Ile Asp Glu Val Val Gly Ala Val Lys
 130 135 140
 Ala Ile Gln Gln Gln Thr Asp Tyr Asp Lys Tyr Cys Leu Met Gly Phe
 145 150 155 160
 Ser Leu Gly Gly Asn Phe Ala Leu Arg Val Ala Val Arg Glu Gln His
 165 170 175
 Leu Ala Lys Pro Leu Ala Gly Val Leu Ala Val Cys Pro Val Leu Asp
 180 185 190
 Pro Ala His Thr Met Met Ala Leu Asn Arg Gly Ala Phe Phe Tyr Gly
 195 200 205
 Arg Tyr Phe Ala His Lys Trp Lys Arg Ser Leu Thr Ala Lys Leu Ala
 210 215 220
 Ala Phe Pro Asp Tyr Lys Tyr Gly Lys Asp Leu Lys Ser Ile His Thr
 225 230 235 240
 30 Leu Asp Glu Leu Asn Asn Tyr Phe Ile Pro Arg Tyr Thr Gly Phe Asn
 245 250 255
 Ser Val Ser Glu Tyr Phe Lys Ser Tyr Thr Leu Thr Gly Gln Lys Leu
 260 265 270
 Ala Phe Leu Asn Cys Pro Ser Tyr Ile Leu Ala Ala Gly Asp Asp Pro
 275 280 285
 35 Ile Ile Pro Ala Ser Asp Phe Gln Lys Ile Ala Lys Pro Ala Asn Leu
 290 295 300
 His Ile Thr Val Thr Gln Gln Gly Ser His Cys Ala Tyr Leu Glu Asn
 305 310 315 320
 Leu His Lys Pro Ser Ala Ala Asp Lys Tyr Ala Val Lys Leu Phe Gly
 325 330 335
 Ala Cys

(2) INFORMATION FOR SEQ ID NO:41:

45 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 311 amino acids
- (B) TYPE: amino acid
- 50 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

55 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Met Leu Asp Met Pro Ile Asp Pro Val Tyr Tyr Gln Leu Ala Glu Tyr
 1 5 10 15
 Phe Asp Ser Leu Pro Lys Phe Asp Gln Phe Ser Ser Ala Arg Glu Tyr
 20 25 30
 Arg Glu Ala Ile Asn Arg Ile Tyr Glu Glu Arg Asn Arg Gln Leu Ser
 35 40 45
 Gln His Glu Arg Val Glu Arg Val Glu Asp Arg Thr Ile Lys Gly Arg
 50 55 60
 Asn Gly Asp Ile Arg Val Arg Val Tyr Gln Gln Lys Pro Asp Ser Pro
 65 70 75 80
 Val Leu Val Tyr Tyr His Gly Gly Phe Val Ile Cys Ser Ile Glu
 85 90 95
 Ser His Asp Ala Leu Cys Arg Arg Ile Ala Arg Leu Ser Asn Ser Thr
 100 105 110
 Val Val Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Ala
 115 120 125
 Ala Val Tyr Asp Cys Tyr Asp Ala Thr Lys Trp Val Ala Glu Asn Ala
 130 135 140
 Glu Glu Leu Arg Ile Asp Pro Ser Lys Ile Phe Val Gly Gly Asp Ser
 145 150 155 160
 Ala Gly Gly Asn Leu Ala Ala Ala Val Ser Ile Met Ala Arg Asp Ser
 165 170 175
 Gly Glu Asp Phe Ile Lys His Gln Ile Leu Ile Tyr Pro Val Val Asn
 180 185 190
 Phe Val Ala Pro Thr Pro Ser Leu Leu Glu Phe Gly Glu Gly Leu Trp
 195 200 205
 Ile Leu Asp Gln Lys Ile Met Ser Trp Phe Ser Glu Gln Tyr Phe Ser
 210 215 220
 Arg Glu Glu Asp Lys Phe Asn Pro Leu Ala Ser Val Ile Phe Ala Asp
 225 230 235 240
 Leu Glu Asn Leu Pro Pro Ala Leu Ile Ile Thr Ala Glu Tyr Asp Pro
 245 250 255
 Leu Arg Asp Glu Gly Glu Val Phe Gly Gln Met Leu Arg Arg Ala Gly
 260 265 270
 Val Glu Ala Ser Ile Val Arg Tyr Arg Gly Val Leu His Gly Phe Ile
 275 280 285
 Asn Tyr Tyr Pro Val Leu Lys Ala Ala Arg Asp Ala Ile Asn Gln Ile
 290 295 300
 Ala Ala Leu Leu Val Phe Asp
 305 310

40

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

45

- (A) LENGTH: 305 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

50

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

55

Met Pro Leu Asp Pro Arg Ile Lys Lys Leu Leu Glu Ser Ala Leu Thr
 1 5 10 15
 Ile Pro Ile Gly Lys Ala Pro Val Glu Glu Val Arg Lys Ile Phe Arg
 20 25 30
 Gln Leu Ala Ser Ala Ala Pro Lys Val Glu Val Gly Lys Val Glu Asp
 35 40 45
 Ile Lys Ile Pro Gly Ser Glu Thr Val Ile Asn Ala Arg Val Tyr Phe
 50 55 60
 Pro Lys Ser Ser Gly Pro Tyr Gly Val Leu Val Tyr Leu His Gly Gly
 65 70 75 80
 Gly Phe Val Ile Gly Asp Val Glu Ser Tyr Asp Pro Leu Cys Arg Ala
 85 90 95
 Ile Thr Asn Ala Cys Asn Cys Val Val Val Ser Val Asp Tyr Arg Leu
 100 105 110
 Ala Pro Glu Tyr Lys Phe Pro Ser Ala Val Ile Asp Ser Phe Asp Ala
 115 120 125
 Thr Asn Trp Val Tyr Asn Asn Leu Asp Lys Phe Asp Gly Lys Met Gly
 130 135 140
 Val Ala Ile Ala Gly Asp Ser Ala Gly Gly Asn Leu Ala Ala Val Val
 145 150 155 160
 Ala Leu Leu Ser Lys Gly Lys Ile Asn Leu Lys Tyr Gln Ile Leu Val
 165 170 175
 Tyr Pro Ala Val Ser Leu Asp Asn Val Ser Arg Ser Met Ile Glu Tyr
 180 185 190
 Ser Asp Gly Phe Phe Leu Thr Arg Glu His Ile Glu Trp Phe Gly Ser
 195 200 205
 Gln Tyr Leu Arg Ser Pro Ala Asp Leu Leu Asp Phe Arg Phe Ser Pro
 210 215 220
 Ile Leu Ala Gln Asp Phe Asn Gly Leu Pro Pro Ala Leu Ile Ile Thr
 225 230 235 240
 Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Ala Tyr Ala Asn Lys
 245 250 255
 Leu Leu Gln Ala Gly Val Ser Val Thr Ser Val Arg Phe Asn Asn Val
 260 265 270
 Ile His Gly Phe Leu Ser Phe Phe Pro Leu Met Glu Gln Gly Arg Asp
 275 280 285
 Ala Ile Gly Leu Ile Gly Ser Val Leu Arg Arg Val Phe Tyr Asp Lys
 290 295 300
 Ile
 305

40

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

45

- (A) LENGTH: 605 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: Genomic DNA

55

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...603

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

EP 0 880 590 B1

	ATG AAG GTT AAA CAC GTT ATT GTT TTA CAT GGC TTA TAT ATG TCT GGC Met Lys Val Lys His Val Ile Val Leu His Gly Leu Tyr Met Ser Gly 1 5 10 15	48
5	TTG GTG ATG CGC CCG TTA TGT TCG CGT CTA GAA GAG TCG GGG GTT AAA Leu Val Met Arg Pro Leu Cys Ser Arg Leu Glu Glu Ser Gly Val Lys 20 25 30	96
10	GTT TTA AAC TTA ACC TAC AAT ACT CGA GAC CCT AAT CGA GAT GCT ATT Val Leu Asn Leu Thr Tyr Asn Thr Arg Asp Pro Asn Arg Asp Ala Ile 35 40 45	144
15	TTT ACG CAA ATA GAT GAG TTT ATT AGC AAT GAG CCT TCT GCT TTA GTG Phe Thr Gln Ile Asp Glu Phe Ile Ser Asn Glu Pro Ser Ala Leu Val 50 55 60	192
20	TGT CAC TCT ATG GGG GGC TTA GTT GCT CGC GCC TAT TTA GAG GCA AAC Cys His Ser Met Gly Gly Leu Val Ala Arg Ala Tyr Leu Glu Ala Asn 65 70 75 80	240
25	TCA GCG CCA AGT CAT CAT GTT GAA AAG GTA ATC ACC TTA GGA ACG CCA Ser Ala Pro Ser His His Val Glu Lys Val Ile Thr Leu Gly Thr Pro 85 90 95	288
30	CAT ACT GGC AGC CAT ATT GCT GAA AAA ATG CAG CAA AAA GGG TTC GAG His Thr Gly Ser His Ile Ala Glu Lys Met Gln Gln Lys Gly Phe Glu 100 105 110	336
35	CTA TTA TTA AAA AAT AGC GTT GAG TTT TTA CTC TCT AAG AAT GGT GAT Leu Leu Leu Lys Asn Ser Val Glu Phe Leu Leu Ser Lys Asn Gly Asp 115 120 125	384
40	TGG CCT TTT AAA GCC AAG CTA TAT AGC ATT GCC GGC GAC TTA CCG ATT Trp Pro Phe Lys Ala Lys Leu Tyr Ser Ile Ala Gly Asp Leu Pro Ile 130 135 140	432
45	GGC TTA ATG CCA CTC ATT GTC AAA GGC AGC CGC TCT GAT GGC ACT GTA Gly Leu Met Pro Leu Ile Val Lys Gly Ser Arg Ser Asp Gly Thr Val 145 150 155 160	480
50	TTG CTA GAT GAA ACC AAG CTA AAG GGT ATG GCT GAA CAC AAG GTG TTT Leu Leu Asp Glu Thr Lys Leu Lys Gly Met Ala Glu His Lys Val Phe 165 170 175	528
	CAT TTA AGC CAT ACA AGT ATG ATT TAC TCT CGC CAA GTC GTT AAT TAT	576
	His Leu Ser His Thr Ser Met Ile Tyr Ser Arg Gln Val Val Asn Tyr 180 185 190	
55	ATT CTT GAG CGC TTG AAC GAG GAC ATT TA Ile Leu Glu Arg Leu Asn Glu Asp Ile 195 200	605

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 779 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

5

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

10

- (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 1...777

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

15

	ATG ATA AAA AAC TTC GAC AGA GAA AAT TCT AGC TTA GTA CTG TCC GGT	48
	Met Ile Lys Asn Phe Asp Arg Glu Asn Ser Ser Leu Val Leu Ser Gly	
	1 5 10 15	
20	GGT GGT GCT CTG GGT ATT GCT CAC TTG GGT GTA CTG CAT GAC CTT GAA	96
	Gly Gly Ala Leu Gly Ile Ala His Leu Gly Val Leu His Asp Leu Glu	
	20 25 30	
25	AAA CAA AAT ATT GTA CCA AAT GAA ATT GTT GGT ACA AGT ATG GGT GGT	144
	Lys Gln Asn Ile Val Pro Asn Glu Ile Val Gly Thr Ser Met Gly Gly	
	35 40 45	
	ATC ATT GGT GCA TCT ATG GCT ATC GGG ATG AAA GAG AAA GAA ATA CTC	192
	Ile Ile Gly Ala Ser Met Ala Ile Gly Met Lys Glu Lys Glu Ile Leu	
	50 55 60	
30	GAA GAA ATC AAA AAC TTT TCC AAT GTC TTC AAC TGG ATA AAA TTC TCT	240
	Glu Glu Ile Lys Asn Phe Ser Asn Val Phe Asn Trp Ile Lys Phe Ser	
	65 70 75 80	
35	TTT TCC GGT AAT TCT GTT GTC GAT AAC GAG AAG ATC GCT AAG ATA TTT	288
	Phe Ser Gly Asn Ser Val Val Asp Asn Glu Lys Ile Ala Lys Ile Phe	
	85 90 95	
40	GAT ACT CTT TTT AAA GAC AGA AAG ATG ACA GAT ACG GTG ATC CCT CTT	336
	Asp Thr Leu Phe Lys Asp Arg Lys Met Thr Asp Thr Val Ile Pro Leu	
	100 105 110	
	AAA CTC ATC GCT ACA AAC TTA CAT AAT GGA CAT AAA AAA GTA TTT ACT	384
	Lys Leu Ile Ala Thr Asn Leu His Asn Gly His Lys Lys Val Phe Thr	
	115 120 125	

45

50

55

EP 0 880 590 B1

	GCT TCG GAT GAT GTA CTG ATC AAA GAT GCA ATA CTC TCA ACA ATG GCA Ala Ser Asp Asp Val Leu Ile Lys Asp Ala Ile Leu Ser Thr Met Ala 130 135 140	432
5	ATA CCC GGT GTA TTT GAA GAA CAT ATT ATT GAT GGT GAA ACC TAT GGC Ile Pro Gly Val Phe Glu Glu His Ile Ile Asp Gly Glu Thr Tyr Gly 145 150 155 160	480
10	GAC GGT TTT CTT TGT GAA AAC CTT GGT GTG AAT GAG GCA ACA TTC AAT Asp Gly Phe Leu Cys Glu Asn Leu Gly Val Asn Glu Ala Thr Phe Asn 165 170 175	528
15	GAT GTT TTA GCT GTA GAT GTC ATG GGT GAG AAC TCT TTT GAA AAA GCA Asp Val Leu Ala Val Asp Val Met Gly Glu Asn Ser Phe Glu Lys Ala 180 185 190	576
20	ATG CCG GAC AAC TTC TTT AAA ACA TCA AAT GTT TTA GAA ATG TTT GAA Met Pro Asp Asn Phe Phe Lys Thr Ser Asn Val Leu Glu Met Phe Glu 195 200 205	624
25	AAA TCA ATG CGA CTT TTT ATT TAC AAC CAG ACA CAG ACA CAT ATT AAA Lys Ser Met Arg Leu Phe Ile Tyr Asn Gln Thr Gln Thr His Ile Lys 210 215 220	672
30	AAT GCA AAT AAA AAT ATT TAT CTT ATT GAA CCC GTT ACC AAA GAG TAT Asn Ala Asn Lys Asn Ile Tyr Leu Ile Glu Pro Val Thr Lys Glu Tyr 225 230 235 240	720
35	AAA ACA TTT CAA TTT CAT AAA CAT AAA GAG ATA CGT GCT TTA GGC TTG Lys Thr Phe Gln Phe His Lys His Lys Glu Ile Arg Ala Leu Gly Leu 245 250 255	768
	GGT TTA CTG TG Gly Leu Leu	779

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 905 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...903

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

55 ATG CCC CTA CAT CCA AAG GTA AAG AAA TTA CTT TCC CAG CTA CCT CCC 48

	Met Pro Leu His Pro Lys Val Lys Lys Leu Leu Ser Gln Leu Pro Pro		
1	5	10	15
5	CAG GAC TTC TCC AGA AAC GTG CAG GAC CTG AGG AAG GCC TGG GAT TTA Gln Asp Phe Ser Arg Asn Val Gln Asp Leu Arg Lys Ala Trp Asp Leu	20	25
		30	
10	CCC TTC TCA GGG AGG AGG GAG ACC CTG AAG AGG GTT GAG GAC CTT GAG Pro Phe Ser Gly Arg Arg Glu Thr Leu Lys Arg Val Glu Asp Leu Glu	35	40
		45	
15	ATA CCC ACT AGG GAC GCA CGA ATC AGG GCC AGG GTC TAC ACC CCC TCA Ile Pro Thr Arg Asp Ala Arg Ile Arg Ala Arg Val Tyr Thr Pro Ser	50	55
		60	
20	AGT AAG GAA AAC TTA CCC GTC CTT GTT TAC TAT CAC GGC GGT GGC TTC Ser Lys Glu Asn Leu Pro Val Leu Val Tyr Tyr His Gly Gly Gly Phe	65	70
		75	80
25	GTC TTC GGT AGC GTT GAC AGC TAC GAC GGC CTC GCA TCC CTT ATT GCC Val Phe Gly Ser Val Asp Ser Tyr Asp Gly Leu Ala Ser Leu Ile Ala	85	90
		95	
30	AAG GAA TCT GGG ATT GCG GTT ATC TCC GTG GAG TAT AGG CTC GCC CCT Lys Glu Ser Gly Ile Ala Val Ile Ser Val Glu Tyr Arg Leu Ala Pro	100	105
		110	
35	GAG CAC AAG TTC CCC ACC GCA GTC AAC GAC TCG TGG GAT GCG CTT CTC Glu His Lys Phe Pro Thr Ala Val Asn Asp Ser Trp Asp Ala Leu Leu	115	120
		125	
40	TGG ATC GCG GAG AAC GGA GGC AAG CTG GGG CTC GAC ACC TCG AGA CTT Trp Ile Ala Glu Asn Gly Gly Lys Leu Gly Leu Asp Thr Ser Arg Leu	130	135
		140	
45	GCC GTG GCT GGG GAT AGT GCT GGA GGA AAC CTG TCT GCC GTG GTG TCC Ala Val Ala Gly Asp Ser Ala Gly Gly Asn Leu Ser Ala Val Val Ser	145	150
		155	160
50	CTC CTG GAC AGG GAC CAG GGT AAG GGA CTG GTT AGT TAT CAG GTC CTA Leu Leu Asp Arg Asp Gln Gly Lys Gly Leu Val Ser Tyr Gln Val Leu	165	170
		175	
55	ATC TAC CCA GCA GTG AAC ATG GTC GAT AAC TCC CCA TCC GTC AGG GAG Ile Tyr Pro Ala Val Asn Met Val Asp Asn Ser Pro Ser Val Arg Glu	180	185
		190	
60	TAC GGC GAG GGA TAC TTC CTC ACC AGG TCC ATG ATG AAC TGG TTC GGG Tyr Gly Glu Gly Tyr Phe Leu Thr Arg Ser Met Met Asn Trp Phe Gly	195	200
		205	
65	ACC ATG TAC TTC TCC TCT GGA AGG GAA GCG GTA TCC CCC TAC GCC TCT Thr Met Tyr Phe Ser Ser Gly Arg Glu Ala Val Ser Pro Tyr Ala Ser	210	215
		220	

EP 0 880 590 B1

5	CCA GCC TTG GCT GAC CTA CAT AAC CTC CCA CCC TCA CTG GTG ATC ACT Pro Ala Leu Ala Asp Leu His Asn Leu Pro Pro Ser Leu Val Ile Thr 225 230 235 240	720
	GCA GAG TAT GAT CCC CTA AGG GAT CAG GGA GAG ACC TAC TCT CAC TCC Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Thr Tyr Ser His Ser 245 250 255	768
10	CTA AAC GAG GCT GGA AAC GTA TCA ACC TTG GTT AGA TAT CAA GGA ATG Leu Asn Glu Ala Gly Asn Val Ser Thr Leu Val Arg Tyr Gln Gly Met 260 265 270	816
15	ATT CAC GGC TTC CTG TCC TTC TAC GAG TGG ATA ACT GCC GGT AAA CTA Ile His Gly Phe Leu Ser Phe Tyr Glu Trp Ile Thr Ala Gly Lys Leu 275 280 285	864
20	GCC ATT CAC CAC ATT GCT GGG GTT CTG AGA TCT GTC CTT TA Ala Ile His His Ile Ala Gly Val Leu Arg Ser Val Leu 290 295 300	905

(2) INFORMATION FOR SEQ ID NO:46:

- 25 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 978 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- 30 (ii) MOLECULE TYPE: Genomic DNA

- 35 (ix) FEATURE:
- (A) NAME/KEY: Coding Sequence
 - (B) LOCATION: 1...975
- 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

40	GTG GCC TTC TTC GAT ATG CCC CTT GAG GAA CTG AAA AAG TAC CGG CCT Val Ala Phe Phe Asp Met Pro Leu Glu Leu Lys Lys Tyr Arg Pro 1 5 10 15	48
45	GAA AGG TAC GAG GAG AAA GAT TTC GAT GAG TTC TGG AGG GAA ACA CTT Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Arg Glu Thr Leu 20 25 30	96
50	AAA GAA AGC GAA GGA TTC CCT CTG GAT CCC GTC TTT GAA AAG GTG GAC Lys Glu Ser Glu Gly Phe Pro Leu Asp Pro Val Phe Glu Lys Val Asp 35 40 45	144
55	TTT CAT CTC AAA ACG GTT GAA ACG TAC GAT GTT ACT TTC TCT GGA TAC Phe His Leu Lys Thr Val Glu Thr Tyr Asp Val Thr Phe Ser Gly Tyr 50 55 60	192

EP 0 880 590 B1

	AGG GGG CAG AGA ATA AAG GGC TGG CTT CTT GTT CCG AAG TTG GCG GAA Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Ala Glu 65 70 75 80	240
5	GAA AAG CTT CCA TGC GTC GTG CAG TAC ATA GGT TAC AAT GGT GGA AGG Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg 85 90 95	288
10	GGT TTT CCA CAC GAC TGG CTG TTC TGG CCG TCA ATG GGT TAC ATC TGT Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys 100 105 110	336
15	TTT GTC ATG GAC ACC AGG GGG CAG GGA AGC GGC TGG ATG AAG GGA GAC Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Met Lys Gly Asp 115 120 125	384
	ACA CCG GAT TAC CCT GAG GGT CCA GTC GAT CCA CAG TAC CCC GGA TTC Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe 130 135 140	432
20	ATG ACG AGG GGC ATT CTG GAT CCG GGA ACC TAT TAC TAC AGG CGA GTC Met Thr Arg Gly Ile Leu Asp Pro Gly Thr Tyr Tyr Tyr Arg Arg Val 145 150 155 160	480
25	TTC GTG GAT GCG GTC AGG GCG GTG GAA GCA GCC ATT TCC TTC CCG AGA Phe Val Asp Ala Val Arg Ala Val Glu Ala Ala Ile Ser Phe Pro Arg 165 170 175	528
30	GTG GAT TCC AGG AAG GTG GTG GTG GAA GCA GCC AGT CAG GGT GGG GGA Val Asp Ser Arg Lys Val Val Ala Gly Gly Ser Gln Gly Gly Gly 180 185 190	576
	ATC CCC CTT GCG GTG AGT GCC CTG TCG AAC AGG GTG AAG GCT CTG CTC Ile Pro Leu Ala Val Ser Ala Leu Ser Asn Arg Val Lys Ala Leu Leu 195 200 205	624
35	TGC GAT GTG CCG TTT CTG TGC CAC TTC AGA AGG GCC GTG CAA CTT GTC Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val 210 215 220	672
40	GAC ACA CAC CCA TAC GTG GAG ATC ACC AAC TTC CTC AAA ACC CAC AGG Asp Thr His Pro Tyr Val Glu Ile Thr Asn Phe Leu Lys Thr His Arg 225 230 235 240	720
	GAC AAA GAG GAG ATT GTT TTC AGA ACA CTT TCC TAC TTC GAT GGT GTG Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val 245 250 255	768
45	AAC TTT GCA GCA AGG GCA AAG GTG CCC GCC CTG TTT TCC GTT GGG CTC Asn Phe Ala Ala Arg Ala Lys Val Pro Ala Leu Phe Ser Val Gly Leu 260 265 270	816
50	ATG GAC ACC ATC TGT CCT CCC TCG ACG GTC TTC GCC GCT TAC AAC CAC Met Asp Thr Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn His 275 280 285	864

EP 0 880 590 B1

(2) INFORMATION FOR SEQ ID NO:47:

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 897 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

25 (ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...894

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

	ATG CGC ACC CTC TCC TTC GGT CCG ATG ACC ACA GGG GGA AGC ATT CAC Met Arg Thr Leu Ser Phe Gly Pro Met Thr Thr Gly Gly Ser Ile His 1 5 10 15	48
35	ATG GCG ACC ATG GAC GTG ATG CGC GGG CCG GGG ATG CAG CGG CTG TCA Met Ala Thr Met Asp Val Met Arg Gly Pro Gly Met Gln Arg Leu Ser 20 25 30	96
40	CAG GGC GCC AGG GAG GCC GCG AAC CAC CCC TGG GCG AAG CGA CTG GGC Gln Gly Ala Arg Glu Ala Ala Asn His Pro Trp Ala Lys Arg Leu Gly 35 40 45	144
45	CGC ATG GGC TAC GCG GCC AAG GGC GCC GTG TAC GCC ATC ATC GGC GTG Arg Met Gly Tyr Ala Ala Lys Gly Ala Val Tyr Ala Ile Ile Gly Val 50 55 60	192
50	CTC GCG CTG AAG CTC GCG GCG GGC GAG GGC GGC CGG ACC ACG GAC AGC Leu Ala Leu Lys Leu Ala Ala Gly Glu Gly Gly Arg Thr Thr Asp Ser 65 70 75 80	240
55	CAC GGC GCG GTG AAC ACC GTG GCG CAC GGG CCC TTC GGC GTC GCG CTG His Gly Ala Val Asn Thr Val Ala His Gly Pro Phe Gly Val Ala Leu 85 90 95	288
60	CTG GCG GTG CTG GTG GTG GGC CTG CTG GGC TAC GTG GTC TGG AGG TTC Leu Ala Val Leu Val Val Gly Leu Leu Gly Tyr Val Val Trp Arg Phe 100 105 110	336

EP 0 880 590 B1

	GCC CAG GCC TTC GTG GAC ACG GAG GAC AAG GGC TCC GAC GCG AAG GGA Ala Gln Ala Phe Val Asp Thr Glu Asp Lys Gly Ser Asp Ala Lys Gly 115 120 125	384
5	ATC GCC ACG CGC GCC ATG TAC TTC CTC AGC GGC TGC ATC TAC GCG TCG Ile Ala Thr Arg Ala Met Tyr Phe Leu Ser Gly Cys Ile Tyr Ala Ser 130 135 140	432
10	CTG GCC TTC TTC GCC GCG CAG TCC CTG GTG GGC GCC GCG CAC GGC CGG Leu Ala Phe Phe Ala Ala Gln Ser Leu Val Gly Ala Ala His Gly Arg 145 150 155 160	480
15	AGC AAG GGG ACG CAG GGC TGG ACG GCC ACG CTG ATG GAG CAG CCC TTT Ser Lys Gly Thr Gln Gly Trp Thr Ala Thr Leu Met Glu Gln Pro Phe 165 170 175	528
20	GGC CGC GTG CTG GTG GCG CTG GTG GGG CTG GGC ATC GTG GGC TTC GCG Gly Arg Val Leu Val Ala Leu Val Gly Leu Gly Ile Val Gly Phe Ala 180 185 190	576
25	CTG AAG CAG TTC CAC ACC GCG TGG AAG GCG AAG TTC CGG GAG AAG CTC Leu Lys Gln Phe His Thr Ala Trp Lys Ala Lys Phe Arg Glu Lys Leu 195 200 205	624
30	ACC CTC ACC GGA CTG GCT GCC CGG AAG CAG CAC CAC ATC GAG CGC ATG Thr Leu Thr Gly Leu Ala Ala Arg Lys Gln His His Ile Glu Arg Met 210 215 220	672
35	TGC CAG TTC GGC ATC GCC GCG CGC GGC GTG GTG TTC GCC GTC ATC GGC Cys Gln Phe Gly Ile Ala Ala Arg Gly Val Val Phe Ala Val Ile Gly 225 230 235 240	720
40	GGC TTC CTC GTC CGC TCC GCC GTG GAC GCG AAC CCC GGC GAG GCC AAG Gly Phe Leu Val Arg Ser Ala Val Asp Ala Asn Pro Gly Glu Ala Lys 245 250 255	768
45	GGC CTG GGA GAG GCC CTG GCC GTC GCG AGG CAG CCG TCC GGC GAC Gly Leu Gly Ala Leu Ala Val Val Ala Arg Gln Pro Ser Gly Asp 260 265 270	816
50	GTG CTC CTG GGG GTG GTG GCG GCG GGC CTG GTG GCC TAC GCC GCC TAC Val Leu Leu Gly Val Val Ala Ala Gly Leu Val Ala Tyr Ala Ala Tyr 275 280 285	864
	CTG TTC CTC CAG GCG CGC TAC CGC GAA CTC TAG Leu Phe Leu Gln Ala Arg Tyr Arg Glu Leu 290 295	897

(2) INFORMATION FOR SEQ ID NO:48:

50 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 914 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

55 (ii) MOLECULE TYPE: Genomic DNA

EP 0 880 590 B1

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...912

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

10	ATG AGC AAA TTC GCA ATA CTC TGG GCG TTG ATA ACG GCA TAC CTG CCG Met Ser Lys Phe Ala Ile Leu Trp Ala Leu Ile Thr Ala Tyr Leu Pro 1 5 10 15	48
15	GAA CCT GTG ATG AAA CTG GTA TAT TTA GGG CGG CGC GAA ACG CTT GGG Glu Pro Val Met Lys Leu Val Tyr Leu Gly Arg Arg Glu Thr Leu Gly 20 25 30	96
20	GCA CGG ACG CTT GAC GTT AAA GCC CAA GCT GTC GGG CGG CTG GCC AAT Ala Arg Thr Leu Asp Val Lys Ala Gln Ala Val Gly Arg Leu Ala Asn 35 40 45	144
25	GCA ACA AGA CCT GTC GGG GTG ATT CCG ACG GTC GAG GAA AGC CGG AAG Ala Thr Arg Pro Val Gly Val Ile Pro Thr Val Glu Glu Ser Arg Lys 50 55 60	192
30	ATG ACG GAT AAA GCC GTT AGC CTT TTT GAT CAG CCC GCC CCC GAA TTA Met Thr Asp Lys Ala Val Ser Leu Phe Asp Gln Pro Ala Pro Glu Leu 65 70 75 80	240
35	TTC CGT AAA AAA GAC ATT CAG ATT GAC GGG GCT GAA GGG CCT ATT GAT Phe Arg Lys Lys Asp Ile Gln Ile Asp Gly Ala Glu Gly Pro Ile Asp 85 90 95	288
40	GCC CGT ATT TAC AGC GGC CCT GCA AAA CAT CGC CCR CGR CCA ATW CTA Ala Arg Ile Tyr Ser Gly Pro Ala Lys His Arg Xaa Xaa Pro Xaa Leu 100 105 110	336
45	GTG TAT TTT CAC GGC GGT GGC TGG GTT CAG GGC AAT CTG GAC AGC CAT Val Tyr Phe His Gly Gly Trp Val Gln Gly Asn Leu Asp Ser His 115 120 125	384
50	GAC GGG GTT TGC GGC AAG CTG GCA AAA TGG GCG AAC TGC ATT GTT ATC Asp Gly Val Cys Gly Lys Leu Ala Lys Trp Ala Asn Cys Ile Val Ile 130 135 140	432
55	TCG GTC GAT TAT CGT CTA GCG CCC GAA CAC AAA TTT CCT TGT GCG CCG Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Cys Ala Pro 145 150 155 160	480
60	CTT GAT GCG ATT GCG GCC TAT AAA TGG GTG CGC GCC AAC GCA ACA AAC Leu Asp Ala Ile Ala Ala Tyr Lys Trp Val Arg Ala Asn Ala Thr Asn 165 170 175	528
65	CTT GGC GGC GAT CCT GAA CGT ATC GGC GTT GGC GGC GAT AGC GCA GGG	576

55

EP 0 880 590 B1

	Leu Gly Gly Asp Pro Glu Arg Ile Gly Val Gly Gly Asp Ser Ala Gly	
	180 185 190	
5	GGC AAT CTT GCC GCC GTT GTC TGC CAA CAA ACC GCC ATG AAC GGC GAG Gly Asn Leu Ala Ala Val Val Cys Gln Gln Thr Ala Met Asn Gly Glu	624
	195 200 205	
10	CGC ACA CCA GAT CTG CAA GTC CTG ATC TAT CCG GCG CTG GAT GCA CGC Arg Thr Pro Asp Leu Gln Val Leu Ile Tyr Pro Ala Leu Asp Ala Arg	672
	210 215 220	
15	ATG ATC TCG ACC TCG ATG GAG GAA TTG CGT GAT GCC TAC ATC TTG CCG Met Ile Ser Thr Ser Met Glu Glu Leu Arg Asp Ala Tyr Ile Leu Pro	720
	225 230 235 240	
	AAA TCC AGA ATG GAG TAT TTC CTC GGC CTA TAT ACG CGT GGC CCT GAC Lys Ser Arg Met Glu Tyr Phe Leu Gly Leu Tyr Thr Arg Gly Pro Asp	768
	245 250 255	
20	GAT ATC GAG GAC CTT AGG ATG TCG CCA ATT CTC AGG GAT ACC GTC GCG Asp Ile Glu Asp Leu Arg Met Ser Pro Ile Leu Arg Asp Thr Val Ala	816
	260 265 270	
25	GAT CAA CCC CAA GCC TGC ATT GTC ACC TGT GGG TTT GAC CCT GCG CGA Asp Gln Pro Gln Ala Cys Ile Val Thr Cys Gly Phe Asp Pro Ala Arg	864
	275 280 285	
	CGA CGG GAA CAC CTA CGC CGA ACG CTT AAT TGC CGA GGG GAT AGA CGT Arg Arg Glu His Leu Arg Arg Thr Leu Asn Cys Arg Gly Asp Arg Arg	912
30	290 295 300	
	TA	914

(2) INFORMATION FOR SEQ ID NO:49:

35	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 926 base pairs	
	(B) TYPE: nucleic acid	
40	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: Genomic DNA	
45	(ix) FEATURE:	
	(A) NAME/KEY: Coding Sequence	
	(B) LOCATION: 1...924	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:	
	GTG AGC ATT CGT CTG CGA CTG TTA AAC TGG TTT TTG AAT ACC TTT GAA Val Ser Ile Arg Leu Arg Leu Leu Asn Trp Phe Leu Asn Thr Phe Glu	48
55	1 5 10 15	
	AAA CCA AAA CTG GCC GCG GCC AAA ACG CCG GAT GAT TTG CGA AAA TCG Lys Pro Lys Leu Ala Ala Lys Thr Pro Asp Asp Leu Arg Lys Ser	96

EP 0 880 590 B1

	20	25	30	
5	TTT GAA TTA AAG GCG AGG TTT TTG TTT CCG GCG CCA CGT AAA ACA AGG Phe Glu Leu Lys Ala Arg Phe Leu Phe Pro Ala Pro Arg Lys Thr Arg 35 40 45			144
10	TTT AGT CAT GAT GTA TTG CAG TCA GGC ATC GGG TCG GTA AAT GCC CAG Phe Ser His Asp Val Leu Gln Ser Gly Ile Gly Ser Val Asn Ala Gln 50 55 60			192
15	TGG GCG AAA TCC AAA TCT GCA TCT GAT GAC AGG GTA ATC CTG TAT TTT Trp Ala Lys Ser Lys Ser Ala Ser Asp Asp Arg Val Ile Leu Tyr Phe 65 70 75 80			240
20	CAT GGG GGA GGG TAT GTT TTT GGG TCA CCA AAA ACG CAC CGT GCA ATG His Gly Gly Tyr Val Phe Gly Ser Pro Lys Thr His Arg Ala Met 85 90 95			288
25	TTG GCG CGC TTG TCG GCA ATG ACA GGT CTT TCT GCG TGC CTT CCA GAT Leu Ala Arg Leu Ser Ala Met Thr Gly Leu Ser Ala Cys Leu Pro Asp 100 105 110			336
30	TAT AGG TTG GCA CCA GAG CAC CCA TTT CCA GCC GCG ATC GAA GAT GCA Tyr Arg Leu Ala Pro Glu His Pro Phe Pro Ala Ala Ile Glu Asp Ala 115 120 125			384
35	GTT TTA TCG TAT AAA TGT TTA CTA GAG CGA GCA ATC GAG CCC CAA AAT Val Leu Ser Tyr Lys Cys Leu Leu Glu Arg Ala Ile Glu Pro Gln Asn 130 135 140			432
40	ATT ATA CTG GGG GGG GAC AGT GCT GGT GGC GGT TTG GTT CTT GCT TTG Ile Ile Leu Gly Gly Asp Ser Ala Gly Gly Leu Val Leu Ala Leu 145 150 155 160			480
45	CTT GCA GAA ATC AAG GCC CAA TCC TTG CCC AAA CCT GCT GGC GTT TTT Leu Ala Glu Ile Lys Ala Gln Ser Leu Pro Lys Pro Ala Gly Val Phe 165 170 175			528
50	GCC TTG TCG CCT TTG GAT TTA TCA TTT TCG GGC CTT TCG TTT TCT Ala Leu Ser Pro Leu Val Asp Leu Ser Phe Ser Gly Leu Ser Phe Ser 180 185 190			576
55	AAA AAT GCC CAA ACC GAT GTG ATG TTG CCC GCA TCA CGG GCT GCG GAT Lys Asn Ala Gln Thr Asp Val Met Leu Pro Ala Ser Arg Ala Ala Asp 195 200 205			624
60	ATG GCG ACC TTG TAT TTG GAT GGG GCC GAT GCA GAT GAT CCA CGT GCA Met Ala Thr Leu Tyr Leu Asp Gly Ala Asp Ala Asp Asp Pro Arg Ala 210 215 220			672
65	TCG CCG CTG CAG GCG GAT TTT TCT GGC ATG CCG CCT GTA TTT CTG ACA Ser Pro Leu Gln Ala Asp Phe Ser Gly Met Pro Pro Val Phe Leu Thr 225 230 235 240			720
70	GCA AGT GAC ACT GAA ATC CTG TTG GAT GAT TGC CTG CGG ATG GCG GAT			768

EP 0 880 590 B1

	Ala Ser Asp Ser Glu Ile Leu Leu Asp Asp Cys Leu Arg Met Ala Asp		
	245	250	255
5	CAC TTG CGT GCG CAA GGT GTC GTT GTG ACA GAC CGG ATT GTT GAA AAC		816
	His Leu Arg Ala Gln Gly Val Val Val Thr Asp Arg Ile Val Glu Asn		
	260	265	270
10	CAT CCA CAT GTT TGG CAT ATT TTT CAA CGC CTT CTA CCC GAA GCA GAT		864
	His Pro His Val Trp His Ile Phe Gln Arg Leu Leu Pro Glu Ala Asp		
	275	280	285
15	CAG GGG CTG CGG GCG ATT GCC GCG TGG ATT AAA CCT CTT TTA TCA GGT		912
	Gln Gly Leu Arg Ala Ile Ala Ala Trp Ile Lys Pro Leu Leu Ser Gly		
	290	295	300
20	TCA AAC GAA AGC TA		926
	Ser Asn Glu Ser		
	305		

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 713 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- 35 (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...711

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

40	ATG CTT ACA TTT AAT GTT TTA TAT GGT ATG ATG AAA CAA AAA CTA GCA		48
	Met Leu Thr Phe Asn Val Leu Tyr Gly Met Met Lys Gln Lys Leu Ala		
	1 5 10 15		
45	GCA ATT CTC ATG TTT TTA GGG CTA TCA GCA GCA GAG GCT CAA GAC TGG		96
	Ala Ile Leu Met Phe Leu Gly Leu Ser Ala Ala Glu Ala Gln Asp Trp		
	20 25 30		
50	CCT GAC CTA CAG AAA TAT CGT AGT GCT AAT AAA GAA GCC AAA TTA CTT		144
	Pro Asp Leu Gln Lys Tyr Arg Ser Ala Asn Lys Glu Ala Lys Leu Leu		
	35 40 45		
55	CCA AAG GAA AAC CGG AAG GTG GTT TTT ATG GGC AAC TCC ATT ACA GAA		192
	Pro Lys Glu Asn Arg Lys Val Val Phe Met Gly Asn Ser Ile Thr Glu		
	50 55 60		
60	GCC TGG ATT AGT CAG CGA CCT GAG TTT TTT AGT GAA AAT GGG TTT ATC		240
	Ala Trp Ile Ser Gln Arg Pro Glu Phe Phe Ser Glu Asn Gly Phe Ile		
	65 70 75 80		

	GGT CGA GGC ATC AGT GGC CAG ACA ACC CCT CAG ATG TTG TTG AGA TTC Gly Arg Gly Ile Ser Gly Gln Thr Thr Pro Gln Met Leu Leu Arg Phe 85 90 95	288
5	CGA CAG GAT GTG ATA GAC CTG CAG CCA AAG GCT GTA GTG ATA CTA GCT Arg Gln Asp Val Ile Asp Leu Gln Pro Lys Ala Val Val Ile Leu Ala 100 105 110	336
10	GGT ACC AAT GAC GTA GCT CAA AAT ACC GGG CCG ATG ACC ATT GAG GAA Gly Thr Asn Asp Val Ala Gln Asn Thr Gly Pro Met Thr Ile Glu Glu 115 120 125	384
15	TCG CTT GCT AAC ATT AAG TCT ATG GTG GAG CTG GCG CAA GCC AAT GGG Ser Leu Ala Asn Ile Lys Ser Met Val Glu Leu Ala Gln Ala Asn Gly 130 135 140	432
20	ATC ACG CCT GTT TTG TGT ACC GTG CTG CCT GCA GAT CGT TTC AGC TGG Ile Thr Pro Val Leu Cys Thr Val Leu Pro Ala Asp Arg Phe Ser Trp 145 150 155 160	480
25	CGA CCT GAG CTT ACA CCC GCA GAA ACT ATC ATT GCC CTC AAT CAG CTC Arg Pro Glu Leu Thr Pro Ala Glu Thr Ile Ile Ala Leu Asn Gln Leu 165 170 175	528
30	ATT AAG CAA TAT GCC GAG GCA CAG GGC CTG GCC CTG GTG GAT TAT CAT Ile Lys Gln Tyr Ala Glu Ala Gln Gly Leu Ala Leu Val Asp Tyr His 180 185 190	576
35	GCT GCA CTC ACC AAT AAA GGT GGA GGA CTT CCG GTG AAA TAC GGA GAA Ala Ala Leu Thr Asn Lys Gly Gly Leu Pro Val Lys Tyr Gly Glu 195 200 205	624
40	GAT GGT GTG CAT CCA AAT GTA GCA GGC TAT CAG GTG ATG GAA AAC ATT Asp Gly Val His Pro Asn Val Ala Gly Tyr Gln Val Met Glu Asn Ile 210 215 220	672
45	GTT TTA CCG GTC ATT TCC AGC GAG TTG GCA AAG CTG AAG TA Val Leu Pro Val Ile Ser Ser Glu Leu Ala Lys Leu Lys 225 230 235	713

40 (2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 978 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...975

55

EP 0 880 590 B1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

5	ATG GCC TTC TTC GAT TTA CCA CTC GAA GAA CTG AAG AAA TAT CGT CCA Met Ala Phe Phe Asp Leu Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro 1 5 10 15	48
10	GAG CGG TAC GAA GAG AAA GAC TTC GAT GAG TTC TGG GAA GAG ACA CTC Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Glu Glu Thr Leu 20 25 30	96
15	GCA GAG AGC GAA AAG TTC CCC TTA GAC CCC GTC TTC GAG AGG ATG GAG Ala Glu Ser Glu Lys Phe Pro Leu Asp Pro Val Phe Glu Arg Met Glu 35 40 45	144
20	TCT CAC CTC AAA ACA GTC GAA GCG TAC GAT GTC ACC TTC TCC GGA TAC Ser His Leu Lys Thr Val Glu Ala Tyr Asp Val Thr Phe Ser Gly Tyr 50 55 60	192
25	AGG GGA CAG AGG ATC AAA GGG TGG CTC CTT GTT CCA AAA CTG GAA GAA Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Glu Glu 65 70 75 80	240
30	GAA AAA CTT CCC TGC GTT GTG CAG TAC ATA GGA TAC AAC GGT GGA AGA Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg 85 90 95	288
35	GGA TTC CCT CAC GAC TGG CTG TTC TGG CCT TCT ATG GGT TAC ATA TGT Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys 100 105 110	336
40	TTC GTC ATG GAT ACT CGA GGT CAG GGA AGC GGC TGG CTG AAA GGA GAC Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Leu Lys Gly Asp 115 120 125	384
45	ACA CCG GAT TAC CCT GAG GGT CCC GTT GAC CCT CAG TAT CCA GGA TTC Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe 130 135 140	432
50	ATG ACA AGA GGA ATA CTG GAT CCC AGA ACT TAC TAC TAC AGA CGA GTC Met Thr Arg Gly Ile Leu Asp Pro Arg Thr Tyr Tyr Arg Arg Val 145 150 155 160	480
55	TTC ACG GAC GCT GTC AGA GCC GTT GAA GCT GCT GCT TCT TTT CCT CAG Phe Thr Asp Ala Val Arg Ala Val Glu Ala Ala Ala Ser Phe Pro Gln 165 170 175	528
60	GTA GAT CAA GAA AGA ATC GTG ATA GCT GGA GGC AGT CAG GGT GGC GGA Val Asp Gln Glu Arg Ile Val Ile Ala Gly Gly Ser Gln Gly Gly Gly 180 185 190	576
65	ATA GCC CTT GCG GTG AGC GCT CTC TCA AAG AAA GCA AAG GCT CTT CTG Ile Ala Leu Ala Val Ser Ala Leu Ser Lys Lys Ala Lys Ala Leu Leu 195 200 205	624
70	TGC GAT GTG CCG TTT CTG TGT CAC TTC AGA AGA GCA GTA CAG CTT GTG	672

EP 0 880 590 B1

	Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val	
	210 215 220	
5	GAT ACG CAT CCA TAC GCG GAG ATC ACG AAC TTT CTA AAG ACC CAC AGA Asp Thr His Pro Tyr Ala Glu Ile Thr Asn Phe Leu Lys Thr His Arg 225 230 235 240	720
	GAC AAG GAA GAA ATC GTG TTC AGG ACT CTT TCC TAT TTC GAT GGA GTG Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val 10 245 250 255	768
	AAC TTC GCA GCC AGA GCG AAG ATC CCT GCG CTG TTT TCT GTG GGT CTC Asn Phe Ala Ala Arg Ala Lys Ile Pro Ala Leu Phe Ser Val Gly Leu 260 265 270	816
15	ATG GAC AAC ATT TGT CCT CCT TCA ACG GTT TTC GCT GCC TAC AAT TAC Met Asp Asn Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn Tyr 275 280 285	864
20	TAC GCT GGA CCG AAG GAA ATC AGA ATC TAT CCG TAC AAC AAC CAC GAG Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu 290 295 300	912
25	GGA GGA GGC TCT TTC CAA GCG GTT GAA CAG GTG AAA TTC TTG AAA AAA Gly Gly Gly Ser Phe Gln Ala Val Glu Gln Val Lys Phe Leu Lys Lys 305 310 315 320	960
	CTA TTT GAG AAA GGC TAA Leu Phe Glu Lys Gly 325	978
30	(2) INFORMATION FOR SEQ ID NO:52:	

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 661 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- 45 (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...630

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

50	TTG AAG TAC TTC AAA GCC CGG CTT GCC GGC ATC ACC TTG CTC GGC CTG Leu Lys Tyr Phe Lys Ala Arg Leu Ala Gly Ile Thr Leu Leu Gly Leu 1 5 10 15	48
55	CTG GCC TGC ACC TCG GCC TCG GCG CAG ACC GAG CCC ATC GTG TTC GTG Leu Ala Cys Thr Ser Ala Ser Ala Gln Thr Glu Pro Ile Val Phe Val 20 25 30	96

5	CAC GGC TAT TCC GGC AGC GCA TCC AAC TGG GAC ACC ATG CTG GGC CGC His Gly Tyr Ser Gly Ser Ala Ser Asn Trp Asp Thr Met Leu Gly Arg 35 40 45	144
10	TTC CGG TCG AAC GGT TAT GCG TCC GGC TCG CTC TAC ACC TTC AAC TAC Phe Arg Ser Asn Gly Tyr Ala Ser Gly Ser Leu Tyr Thr Phe Asn Tyr 50 55 60	192
15	AAC TCG TTG GTC AGC AGC AAC CGC ACC AGC GCC AGC GAG CTG CGC AGC Asn Ser Leu Val Ser Ser Asn Arg Thr Ser Ala Ser Glu Leu Arg Ser 65 70 75 80	240
20	TTC GTC AAC ACC GTG CGT TCG CGC CAC GGC AAC GCC CGC ATC GCG CTG Phe Val Asn Thr Val Arg Ser Arg His Gly Asn Ala Arg Ile Ala Leu 85 90 95	288
25	GTC GCC CAC TCC AAC GGC GGG CTG GTG TCG CGC TGG TAT CGC GCG GAG Val Ala His Ser Asn Gly Gly Leu Val Ser Arg Trp Tyr Arg Ala Glu 100 105 110	336
30	CTG GGC GGC GAA ACG GCC ACC CGC CGC TTC GTG ACG CTG GGC ACG CCG Leu Gly Gly Glu Thr Ala Thr Arg Arg Phe Val Thr Leu Gly Thr Pro 115 120 125	384
35	CAC CGG GGC ACC ACC TGG GCC TAT GCG TGC TAC AGC CCC GCA TGT TTC His Arg Gly Thr Thr Trp Ala Tyr Ala Cys Tyr Ser Pro Ala Cys Phe 130 135 140	432
40	GAG ATG CGC CCC GGC TCC AGC TTG CTG ACC ACG CTG GGC TCG CGT GCC Glu Met Arg Pro Gly Ser Ser Leu Leu Thr Thr Leu Gly Ser Arg Ala 145 150 155 160	480
45	TGC GAC CGC TCG CTG TGG TCG AAC ACC GAC GGC ATC ATC CTG CCG GCG Cys Asp Arg Ser Leu Trp Ser Asn Thr Asp Gly Ile Ile Leu Pro Ala 165 170 175	528
50	TCC AGC GCG CAG TGT GGT GTC AGC ACG CGC ACT GCC GAC GTC AGC CAT Ser Ser Ala Gln Cys Gly Val Ser Thr Arg Thr Ala Asp Val Ser His 180 185 190	576
55	CTC GAC CTG CTG ACC GAC TCT CGC GTG TAC ACG CAG TTG CGC ACG CAG Leu Asp Leu Leu Thr Asp Ser Arg Val Tyr Thr Gln Leu Arg Thr Gln 195 200 205	624
60	TTG CAA TGAGGGTGAC GGTGCACCGA ACGTGCACCT G	661
65	Leu Gln 210	

(2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 201 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

EP 0 880 590 B1

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

5 Met Lys Val Lys His Val Ile Val Leu His Gly Leu Tyr Met Ser Gly
1 5 10 15
Leu Val Met Arg Pro Leu Cys Ser Arg Leu Glu Glu Ser Gly Val Lys
20 25 30
10 Val Leu Asn Leu Thr Tyr Asn Thr Arg Asp Pro Asn Arg Asp Ala Ile
35 40 45
Phe Thr Gln Ile Asp Glu Phe Ile Ser Asn Glu Pro Ser Ala Leu Val
50 55 60
Cys His Ser Met Gly Gly Leu Val Ala Arg Ala Tyr Leu Glu Ala Asn
65 70 75 80
15 Ser Ala Pro Ser His His Val Glu Lys Val Ile Thr Leu Gly Thr Pro
85 90 95
His Thr Gly Ser His Ile Ala Glu Lys Met Gln Gln Lys Gly Phe Glu
100 105 110
Leu Leu Leu Lys Asn Ser Val Glu Phe Leu Leu Ser Lys Asn Gly Asp
115 120 125
20 Trp Pro Phe Lys Ala Lys Leu Tyr Ser Ile Ala Gly Asp Leu Pro Ile
130 135 140
Gly Leu Met Pro Leu Ile Val Lys Gly Ser Arg Ser Asp Gly Thr Val
145 150 155 160
Leu Leu Asp Glu Thr Lys Leu Lys Gly Met Ala Glu His Lys Val Phe
165 170 175
25 His Leu Ser His Thr Ser Met Ile Tyr Ser Arg Gln Val Val Asn Tyr
180 185 190
Ile Leu Glu Arg Leu Asn Glu Asp Ile
195 200

30 (2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 259 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

40 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

45 Met Ile Lys Asn Phe Asp Arg Glu Asn Ser Ser Leu Val Leu Ser Gly
1 5 10 15
Gly Gly Ala Leu Gly Ile Ala His Leu Gly Val Leu His Asp Leu Glu
20 25 30
50 Lys Gln Asn Ile Val Pro Asn Glu Ile Val Gly Thr Ser Met Gly Gly
35 40 45
Ile Ile Gly Ala Ser Met Ala Ile Gly Met Lys Glu Lys Glu Ile Leu
50 55 60
Glu Glu Ile Lys Asn Phe Ser Asn Val Phe Asn Trp Ile Lys Phe Ser

55

65 70 75 80
 Phe Ser Gly Asn Ser Val Val Asp Asn Glu Lys Ile Ala Lys Ile Phe
 85 90 95
 5 Asp Thr Leu Phe Lys Asp Arg Lys Met Thr Asp Thr Val Ile Pro Leu
 100 105 110
 Lys Leu Ile Ala Thr Asn Leu His Asn Gly His Lys Lys Val Phe Thr
 115 120 125
 Ala Ser Asp Asp Val Leu Ile Lys Asp Ala Ile Leu Ser Thr Met Ala
 130 135 140
 Ile Pro Gly Val Phe Glu Glu His Ile Ile Asp Gly Glu Thr Tyr Gly
 145 150 155 160
 Asp Gly Phe Leu Cys Glu Asn Leu Gly Val Asn Glu Ala Thr Phe Asn
 165 170 175
 15 Asp Val Leu Ala Val Asp Val Met Gly Glu Asn Ser Phe Glu Lys Ala
 180 185 190
 Met Pro Asp Asn Phe Phe Lys Thr Ser Asn Val Leu Glu Met Phe Glu
 195 200 205
 20 Lys Ser Met Arg Leu Phe Ile Tyr Asn Gln Thr Gln Thr His Ile Lys
 210 215 220
 Asn Ala Asn Lys Asn Ile Tyr Leu Ile Glu Pro Val Thr Lys Glu Tyr
 225 230 235 240
 Lys Thr Phe Gln Phe His Lys His Lys Glu Ile Arg Ala Leu Gly Leu
 245 250 255
 Gly Leu Leu
 25

(2) INFORMATION FOR SEQ ID NO:55:

30 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 301 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Met Pro Leu His Pro Lys Val Lys Lys Leu Leu Ser Gln Leu Pro Pro
 1 5 10 15
 45 Gln Asp Phe Ser Arg Asn Val Gln Asp Leu Arg Lys Ala Trp Asp Leu
 20 25 30
 Pro Phe Ser Gly Arg Arg Glu Thr Leu Lys Arg Val Glu Asp Leu Glu
 35 40 45
 Ile Pro Thr Arg Asp Ala Arg Ile Arg Ala Arg Val Tyr Thr Pro Ser
 50 55 60
 50 Ser Lys Glu Asn Leu Pro Val Leu Val Tyr Tyr His Gly Gly Gly Phe
 65 70 75 80
 Val Phe Gly Ser Val Asp Ser Tyr Asp Gly Leu Ala Ser Leu Ile Ala
 85 90 95
 55 Lys Glu Ser Gly Ile Ala Val Ile Ser Val Glu Tyr Arg Leu Ala Pro
 100 105 110
 Glu His Lys Phe Pro Thr Ala Val Asn Asp Ser Trp Asp Ala Leu Leu

	115	120	125
	Trp Ile Ala Glu Asn Gly Gly	Lys Leu Gly Leu Asp Thr Ser Arg Leu	
5	130	135	140
	Ala Val Ala Gly Asp Ser Ala Gly Gly Asn Leu Ser Ala Val Val Ser		
	145	150	155
	Leu Leu Asp Arg Asp Gln Gly Lys Gly Leu Val Ser Tyr Gln Val Leu		160
	165	170	175
10	Ile Tyr Pro Ala Val Asn Met Val Asp Asn Ser Pro Ser Val Arg Glu		
	180	185	190
	Tyr Gly Glu Gly Tyr Phe Leu Thr Arg Ser Met Met Asn Trp Phe Gly		
	195	200	205
	Thr Met Tyr Phe Ser Ser Gly Arg Glu Ala Val Ser Pro Tyr Ala Ser		
15	210	215	220
	Pro Ala Leu Ala Asp Leu His Asn Leu Pro Pro Ser Leu Val Ile Thr		
	225	230	235
	Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Thr Tyr Ser His Ser		240
	245	250	255
	Leu Asn Glu Ala Gly Asn Val Ser Thr Leu Val Arg Tyr Gln Gly Met		
20	260	265	270
	Ile His Gly Phe Leu Ser Phe Tyr Glu Trp Ile Thr Ala Gly Lys Leu		
	275	280	285
	Ala Ile His His Ile Ala Gly Val Leu Arg Ser Val Leu		
	290	295	300

25 (2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 325 amino acids
- (B) TYPE: amino acid
- (C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

35 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

40	Val Ala Phe Phe Asp Met Pro Leu Glu Glu Leu Lys Tyr Arg Pro		
	1 5	10	15
	Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Arg Glu Thr Leu		
	20 25	30	
	Lys Glu Ser Glu Gly Phe Pro Leu Asp Pro Val Phe Glu Lys Val Asp		
45	35 40	45	
	Phe His Leu Lys Thr Val Glu Thr Tyr Asp Val Thr Phe Ser Gly Tyr		
	50 55	60	
	Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Ala Glu		
	65 70	75	80
	Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg		
50	85 90	95	
	Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys		
	100 105	110	
	Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Met Lys Gly Asp		
	115 120	125	
55	Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe		

	130	135	140
5	Met Thr Arg Gly Ile Leu Asp Pro Gly Thr Tyr Tyr Tyr Arg Arg Val		
	145	150	155
	Phe Val Asp Ala Val Arg Ala Val Glu Ala Ala Ile Ser Phe Pro Arg		
	165	170	175
	Val Asp Ser Arg Lys Val Val Val Ala Gly Gly Ser Gln Gly Gly Gly		
	180	185	190
10	Ile Pro Leu Ala Val Ser Ala Leu Ser Asn Arg Val Lys Ala Leu Leu		
	195	200	205
	Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val		
	210	215	220
	Asp Thr His Pro Tyr Val Glu Ile Thr Asn Phe Leu Lys Thr His Arg		
15	225	230	235
	Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val		
	245	250	255
	Asn Phe Ala Ala Arg Ala Lys Val Pro Ala Leu Phe Ser Val Gly Leu		
	260	265	270
20	Met Asp Thr Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn His		
	275	280	285
	Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu		
	290	295	300
	Gly Gly Gly Ser Phe Gln Ala Ile Glu Gln Val Lys Phe Leu Lys Arg		
	305	310	315
25	Leu Phe Glu Glu Gly		
		325	

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 298 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

	Met Arg Thr Leu Ser Phe Gly Pro Met Thr Thr Gly Gly Ser Ile His		
	1	5	10
45	Met Ala Thr Met Asp Val Met Arg Gly Pro Gly Met Gln Arg Leu Ser		
	20	25	30
	Gln Gly Ala Arg Glu Ala Ala Asn His Pro Trp Ala Lys Arg Leu Gly		
	35	40	45
	Arg Met Gly Tyr Ala Ala Lys Gly Ala Val Tyr Ala Ile Ile Gly Val		
	50	55	60
50	Leu Ala Leu Lys Leu Ala Ala Gly Glu Gly Gly Arg Thr Thr Asp Ser		
	65	70	75
	His Gly Ala Val Asn Thr Val Ala His Gly Pro Phe Gly Val Ala Leu		
	85	90	95
	Leu Ala Val Leu Val Val Gly Leu Leu Gly Tyr Val Val Trp Arg Phe		
	100	105	110
55	Ala Gln Ala Phe Val Asp Thr Glu Asp Lys Gly Ser Asp Ala Lys Gly		
	115	120	125

Ile Ala Thr Arg Ala Met Tyr Phe Leu Ser Gly Cys Ile Tyr Ala Ser
 130 135 140
 5 Leu Ala Phe Phe Ala Ala Gln Ser Leu Val Gly Ala Ala His Gly Arg
 145 150 155 160
 Ser Lys Gly Thr Gln Gly Trp Thr Ala Thr Leu Met Glu Gln Pro Phe
 165 170 175
 Gly Arg Val Leu Val Ala Leu Val Gly Leu Gly Ile Val Gly Phe Ala
 180 185 190
 10 Leu Lys Gln Phe His Thr Ala Trp Lys Ala Lys Phe Arg Glu Lys Leu
 195 200 205
 Thr Leu Thr Gly Leu Ala Ala Arg Lys Gln His His Ile Glu Arg Met
 210 215 220
 15 Cys Gln Phe Gly Ile Ala Ala Arg Gly Val Val Phe Ala Val Ile Gly
 225 230 235 240
 Gly Phe Leu Val Arg Ser Ala Val Asp Ala Asn Pro Gly Glu Ala Lys
 245 250 255
 Gly Leu Gly Glu Ala Leu Ala Val Val Ala Arg Gln Pro Ser Gly Asp
 260 265 270
 20 Val Leu Leu Gly Val Val Ala Ala Gly Leu Val Ala Tyr Ala Ala Tyr
 275 280 285
 Leu Phe Leu Gln Ala Arg Tyr Arg Glu Leu
 290 295

25

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 304 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 35 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

40

Met Ser Lys Phe Ala Ile Leu Trp Ala Leu Ile Thr Ala Tyr Leu Pro
 1 5 10 15
 Glu Pro Val Met Lys Leu Val Tyr Leu Gly Arg Arg Glu Thr Leu Gly
 20 25 30
 45 Ala Arg Thr Leu Asp Val Lys Ala Gln Ala Val Gly Arg Leu Ala Asn
 35 40 45
 Ala Thr Arg Pro Val Gly Val Ile Pro Thr Val Glu Glu Ser Arg Lys
 50 55 60
 Met Thr Asp Lys Ala Val Ser Leu Phe Asp Gln Pro Ala Pro Glu Leu
 65 70 75 80
 50 Phe Arg Lys Lys Asp Ile Gln Ile Asp Gly Ala Glu Gly Pro Ile Asp
 85 90 95
 Ala Arg Ile Tyr Ser Gly Pro Ala Lys His Arg Xaa Xaa Pro Xaa Leu
 100 105 110
 55 Val Tyr Phe His Gly Gly Trp Val Gln Gly Asn Leu Asp Ser His
 115 120 125

Asp Gly Val Cys Gly Lys Leu Ala Lys Trp Ala Asn Cys Ile Val Ile
 130 135 140
 Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Cys Ala Pro
 145 150 155 160
 Leu Asp Ala Ile Ala Ala Tyr Lys Trp Val Arg Ala Asn Ala Thr Asn
 165 170 175
 Leu Gly Gly Asp Pro Glu Arg Ile Gly Val Gly Gly Asp Ser Ala Gly
 180 185 190
 Gly Asn Leu Ala Ala Val Val Cys Gln Gln Thr Ala Met Asn Gly Glu
 195 200 205
 Arg Thr Pro Asp Leu Gln Val Leu Ile Tyr Pro Ala Leu Asp Ala Arg
 210 215 220
 Met Ile Ser Thr Ser Met Glu Glu Leu Arg Asp Ala Tyr Ile Leu Pro
 225 230 235 240
 Lys Ser Arg Met Glu Tyr Phe Leu Gly Leu Tyr Thr Arg Gly Pro Asp
 245 250 255
 Asp Ile Glu Asp Leu Arg Met Ser Pro Ile Leu Arg Asp Thr Val Ala
 260 265 270
 Asp Gln Pro Gln Ala Cys Ile Val Thr Cys Gly Phe Asp Pro Ala Arg
 275 280 285
 Arg Arg Glu His Leu Arg Arg Thr Leu Asn Cys Arg Gly Asp Arg Arg
 290 295 300

25

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

30

- (A) LENGTH: 308 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

35

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

40

Val Ser Ile Arg Leu Arg Leu Leu Asn Trp Phe Leu Asn Thr Phe Glu
 1 5 10 15
 Lys Pro Lys Leu Ala Ala Ala Lys Thr Pro Asp Asp Leu Arg Lys Ser
 20 25 30
 Phe Glu Leu Lys Ala Arg Phe Leu Phe Pro Ala Pro Arg Lys Thr Arg
 35 40 45
 Phe Ser His Asp Val Leu Gln Ser Gly Ile Gly Ser Val Asn Ala Gln
 50 55 60
 Trp Ala Lys Ser Lys Ser Ala Ser Asp Asp Arg Val Ile Leu Tyr Phe
 65 70 75 80
 His Gly Gly Gly Tyr Val Phe Gly Ser Pro Lys Thr His Arg Ala Met
 85 90 95
 Leu Ala Arg Leu Ser Ala Met Thr Gly Leu Ser Ala Cys Leu Pro Asp
 100 105 110
 Tyr Arg Leu Ala Pro Glu His Pro Phe Pro Ala Ala Ile Glu Asp Ala
 115 120 125
 Val Leu Ser Tyr Lys Cys Leu Leu Glu Arg Ala Ile Glu Pro Gln Asn
 130 135 140

5 Ile Ile Leu Gly Gly Asp Ser Ala Gly Gly Leu Val Leu Ala Leu
 145 150 155 160
 Leu Ala Glu Ile Lys Ala Gln Ser Leu Pro Lys Pro Ala Gly Val Phe
 165 170 175
 Ala Leu Ser Pro Leu Val Asp Leu Ser Phe Ser Gly Leu Ser Phe Ser
 180 185 190
 Lys Asn Ala Gln Thr Asp Val Met Leu Pro Ala Ser Arg Ala Ala Asp
 195 200 205
 10 Met Ala Thr Leu Tyr Leu Asp Gly Ala Asp Ala Asp Pro Arg Ala
 210 215 220
 Ser Pro Leu Gln Ala Asp Phe Ser Gly Met Pro Pro Val Phe Leu Thr
 225 230 235 240
 Ala Ser Asp Ser Glu Ile Leu Leu Asp Asp Cys Leu Arg Met Ala Asp
 245 250 255
 His Leu Arg Ala Gln Gly Val Val Val Thr Asp Arg Ile Val Glu Asn
 260 265 270
 His Pro His Val Trp His Ile Phe Gln Arg Leu Leu Pro Glu Ala Asp
 275 280 285
 20 Gln Gly Leu Arg Ala Ile Ala Ala Trp Ile Lys Pro Leu Leu Ser Gly
 290 295 300
 Ser Asn Glu Ser
 305

25 (2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 237 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

35 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

40 Met Leu Thr Phe Asn Val Leu Tyr Gly Met Met Lys Gln Lys Leu Ala
 1 5 10 15
 Ala Ile Leu Met Phe Leu Gly Leu Ser Ala Ala Glu Ala Gln Asp Trp
 20 25 30
 Pro Asp Leu Gln Lys Tyr Arg Ser Ala Asn Lys Glu Ala Lys Leu Leu
 35 40 45
 Pro Lys Glu Asn Arg Lys Val Val Phe Met Gly Asn Ser Ile Thr Glu
 50 55 60
 Ala Trp Ile Ser Gln Arg Pro Glu Phe Phe Ser Glu Asn Gly Phe Ile
 65 70 75 80
 Gly Arg Gly Ile Ser Gly Gln Thr Thr Pro Gln Met Leu Leu Arg Phe
 85 90 95
 Arg Gln Asp Val Ile Asp Leu Gln Pro Lys Ala Val Val Ile Leu Ala
 100 105 110
 Gly Thr Asn Asp Val Ala Gln Asn Thr Gly Pro Met Thr Ile Glu Glu
 115 120 125
 55 Ser Leu Ala Asn Ile Lys Ser Met Val Glu Leu Ala Gln Ala Asn Gly
 130 135 140

5 Ile Thr Pro Val Leu Cys Thr Val Leu Pro Ala Asp Arg Phe Ser Trp
 145 150 155 160
 Arg Pro Glu Leu Thr Pro Ala Glu Thr Ile Ile Ala Leu Asn Gln Leu
 165 170 175
 Ile Lys Gln Tyr Ala Glu Ala Gln Gly Leu Ala Leu Val Asp Tyr His
 180 185 190
 Ala Ala Leu Thr Asn Lys Gly Gly Leu Pro Val Lys Tyr Gly Glu
 195 200 205
 10 Asp Gly Val His Pro Asn Val Ala Gly Tyr Gln Val Met Glu Asn Ile
 210 215 220
 Val Leu Pro Val Ile Ser Ser Glu Leu Ala Lys Leu Lys
 225 230 235

15 (2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 325 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

25 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

30 Met Ala Phe Phe Asp Leu Pro Leu Glu Glu Leu Lys Lys Tyr Arg Pro
 1 5 10 15
 Glu Arg Tyr Glu Glu Lys Asp Phe Asp Glu Phe Trp Glu Glu Thr Leu
 20 25 30
 Ala Glu Ser Glu Lys Phe Pro Leu Asp Pro Val Phe Glu Arg Met Glu
 35 40 45
 Ser His Leu Lys Thr Val Glu Ala Tyr Asp Val Thr Phe Ser Gly Tyr
 50 55 60
 Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Glu Glu
 65 70 75 80
 40 Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Arg
 85 90 95
 Gly Phe Pro His Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys
 100 105 110
 Phe Val Met Asp Thr Arg Gly Gln Gly Ser Gly Trp Leu Lys Gly Asp
 115 120 125
 45 Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln Tyr Pro Gly Phe
 130 135 140
 Met Thr Arg Gly Ile Leu Asp Pro Arg Thr Tyr Tyr Arg Arg Val
 145 150 155 160
 Phe Thr Asp Ala Val Arg Ala Val Glu Ala Ala Ala Ser Phe Pro Gln
 165 170 175
 50 Val Asp Gln Glu Arg Ile Val Ile Ala Gly Gly Ser Gln Gly Gly
 180 185 190
 Ile Ala Leu Ala Val Ser Ala Leu Ser Lys Lys Ala Lys Ala Leu Leu
 195 200 205
 55 Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala Val Gln Leu Val
 210 215 220
 Asp Thr His Pro Tyr Ala Glu Ile Thr Asn Phe Leu Lys Thr His Arg

225 230 235 240
 5 Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val
 245 250 255
 Asn Phe Ala Ala Arg Ala Lys Ile Pro Ala Leu Phe Ser Val Gly Leu
 260 265 270
 Met Asp Asn Ile Cys Pro Pro Ser Thr Val Phe Ala Ala Tyr Asn Tyr
 275 280 285
 10 Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr Asn Asn His Glu
 290 295 300
 Gly Gly Gly Ser Phe Gln Ala Val Glu Gln Val Lys Phe Leu Lys Lys
 305 310 315 320
 Leu Phe Glu Lys Gly
 325

15

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

20

- (A) LENGTH: 210 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

25

- (ii) MOLECULE TYPE: protein

- (v) FRAGMENT TYPE: internal

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

35

Leu Lys Tyr Phe Lys Ala Arg Leu Ala Gly Ile Thr Leu Leu Gly Leu
 1 5 10 15
 Leu Ala Cys Thr Ser Ala Ser Ala Gln Thr Glu Pro Ile Val Phe Val
 20 25 30
 His Gly Tyr Ser Gly Ser Ala Ser Asn Trp Asp Thr Met Leu Gly Arg
 35 40 45
 Phe Arg Ser Asn Gly Tyr Ala Ser Gly Ser Leu Tyr Thr Phe Asn Tyr
 50 55 60
 40 Asn Ser Leu Val Ser Ser Asn Arg Thr Ser Ala Ser Glu Leu Arg Ser
 65 70 75 80
 Phe Val Asn Thr Val Arg Ser Arg His Gly Asn Ala Arg Ile Ala Leu
 85 90 95
 Val Ala His Ser Asn Gly Gly Leu Val Ser Arg Trp Tyr Arg Ala Glu
 100 105 110
 45 Leu Gly Gly Glu Thr Ala Thr Arg Arg Phe Val Thr Leu Gly Thr Pro
 115 120 125
 His Arg Gly Thr Thr Trp Ala Tyr Ala Cys Tyr Ser Pro Ala Cys Phe
 130 135 140
 Glu Met Arg Pro Gly Ser Ser Leu Leu Thr Thr Leu Gly Ser Arg Ala
 145 150 155 160
 50 Cys Asp Arg Ser Leu Trp Ser Asn Thr Asp Gly Ile Ile Leu Pro Ala
 165 170 175
 Ser Ser Ala Gln Cys Gly Val Ser Thr Arg Thr Ala Asp Val Ser His
 180 185 190
 55 Leu Asp Leu Leu Thr Asp Ser Arg Val Tyr Thr Gln Leu Arg Thr Gln
 195 200 205
 Leu Gln
 210

Claims

1. A polynucleotide comprising a member selected from the group consisting of:
 - 5 a) a polynucleotide comprising a nucleic acid sequence set forth in SEQ ID NO:23;
 - b) a polynucleotide which encodes an enzyme comprising an amino acid sequence set forth in SEQ ID NO:33;
 - c) a polynucleotide comprising at least 15 consecutive bases of a polynucleotide of (a) or (b);
 - d) a polynucleotide having at least 70%, 80%, 90%, 95% or 97% identity to the polynucleotide of (a) or (b);
and encoding an enzyme having esterase activity; and
 - 10 e) a polynucleotide which is complementary to a polynucleotide of any one of (a) to (d).
2. A polynucleotide hybridizing under stringent conditions to:
 - 15 a) a polynucleotide comprising a nucleic acid sequence set forth in SEQ ID NO:23; or
 - b) a polynucleotide which is complementary to a polynucleotide of (a);
and encoding an enzyme having esterase activity.
3. The polynucleotide of claim 1 or 2, wherein the polynucleotide is DNA or RNA.
20
4. A vector comprising the polynucleotide of any one of claims 1 to 3.
5. A host cell comprising the vector of claim 4.
- 25 6. A process for producing a polypeptide comprising expressing from the host cell of claim 5 a polypeptide encoded by said polynucleotide.
7. A process for producing a cell comprising transforming or transfecting the cell with the vector of claim 4 such that
the cell expresses the polypeptide encoded by the polynucleotide contained in the vector.
30
- 35 8. An enzyme comprising a member selected from the group consisting of:
 - (a) an enzyme encoded by the polynucleotide of any one of claims 1(a), (b) or (d);
 - (b) an enzyme comprising an amino acid sequence which is at least 70%, 90% or 95% identical to an amino
acid sequence set forth in SEQ ID NO:33 ;
 - (c) an enzyme which comprises at least 30 amino acids of the enzyme of (a) or (b); and
 - (d) an enzyme obtainable by the process of claim 6.
9. An antibody specifically recognizing the enzyme of claim 8.
40
10. A method for transferring an amino group from an amino acid to an α -keto acid comprising contacting an amino
acid in the presence of an α -keto acid with the enzyme of claim 8.
11. A method for the isolation of the enzyme of claim 8 comprising using the antibody of claim 9.
45
12. Use of the enzyme of claim 8:
 - 50 (a) as ripening starter for cheese in the dairy industry;
 - (b) for lignin removal in the pulp and paper industry;
 - (c) for the synthesis of carbohydrate derivatives; and/or
 - (d) for the conversion of lignocellulosic wastes to fermentable sugars.
13. Use of the polynucleotide of claim 1 or 2 for the identification of similar sequences in other organisms.
55

Patentansprüche

1. Polynukleotid, umfassend ein Mitglied ausgewählt aus der Gruppe bestehend aus:

EP 0 880 590 B1

- (a) einem Polynucleotid, umfassend eine Nucleinsäuresequenz wie in SEQ ID NR:23 dargestellt;
(b) einem Polynucleotid, das ein Enzym codiert, umfassend eine Aminosäuresequenz wie in SEQ ID NR:33 dargestellt;
5 (c) einem Polynucleotid, umfassend mindestens 15 aufeinanderfolgende Basen eines Polynucleotids aus (a) oder (b);
(d) einem Polynucleotid, das mindestens 70%, 80%, 90%, 95% oder 97% Identität mit dem Polynucleotid aus (a) oder (b) hat; und ein Enzym mit Esterase-Aktivität codiert; und
(e) einem Polynucleotid, das zu einem Polynucleotid aus (a) bis (d) komplementär ist.

10 2. Polynucleotid, das unter stringenten Bedingungen an:

- (a) ein Polynucleotid, umfassend eine Nucleinsäuresequenz wie in SEQ ID NR:23 dargestellt; oder
(b) ein Polynucleotid, das zu einem Polynucleotid aus (a) komplementär ist; hybridisiert und ein Enzym mit
15 Esterase-Aktivität codiert.

15 3. Polynucleotid nach Anspruch 1 oder 2, wobei das Polynucleotid DNA oder RNA ist.

20 4. Vektor, umfassend das Polynucleotid nach einem der Ansprüche 1 bis 3.

5. Wirtszelle, umfassend den Vektor nach Anspruch 4.

25 6. Verfahren zur Herstellung eines Polypeptids, umfassend das Exprimieren eines Polypeptids, das von dem Polynucleotid codiert wird, von der Wirtszelle nach Anspruch 5.

7. Verfahren zur Herstellung einer Zelle, umfassend das Transformieren oder Transfizieren der Zelle mit dem Vektor nach Anspruch 4, so dass die Zelle das Polypeptid exprimiert, das von dem im Vektor enthaltenen Polynucleotid codiert wird.

30 8. Enzym, umfassend ein Mitglied ausgewählt aus der Gruppe bestehend aus:

- (a) einem Enzym, das von dem Polynucleotid nach einem der Ansprüche 1 (a), (b) oder (d) codiert wird;
(b) einem Enzym, umfassend eine Aminosäuresequenz, die zu mindestens 70%, 90% oder 95% identisch ist zu einer Aminosäuresequenz wie in SEQ ID NR:33 dargestellt;
35 (c) einem Enzym, das mindestens 30 Aminosäuren des Enzyms aus (a) oder (b) umfasst; und
(d) einem Enzym, erhältlich durch das Verfahren nach Anspruch 6.

9. Antikörper, der spezifisch das Enzym nach Anspruch 8 erkennt.

40 10. Verfahren zum Transferieren einer Aminogruppe von einer Aminosäure zu einer α -Ketosäure, umfassend das Inkontaktbringen einer Aminosäure mit dem Enzym nach Anspruch 8 in Anwesenheit einer α -Ketosäure.

11. Verfahren zum Isolieren des Enzyms nach Anspruch 8, umfassend die Verwendung des Antikörpers nach Anspruch 9.

45 12. Verwendung des Enzyms nach Anspruch 8:

- (a) als Reifestarter für Käse in der Milchindustrie;
(b) zur Ligninenentfernung in der Zellstoffindustrie;
(c) zur Synthese von Kohlenhydratderivaten; und/oder
50 (d) zur Umwandlung von Lignozelluloseabfällen zu fermentierbaren Zuckern.

13. Verwendung des Polynucleotids nach Anspruch 1 oder 2 zur Identifizierung von ähnlichen Sequenzen in anderen Organismen.

55

Revendications

1. Polynucléotide comprenant un membre choisi dans le groupe constitué de :

- a) un polynucléotide comprenant une séquence d'acide nucléique représentée dans SEQ ID NO : 23 ;
b) un polynucléotide qui code une enzyme comprenant une séquence d'acides aminés représentée dans SEQ ID NO : 33 ;
c) un polynucléotide comprenant au moins 15 bases consécutives d'un polynucléotide de a) ou b) ;
d) un polynucléotide ayant au moins 70%, 80%, 90%, 95% ou 97% d'identité avec le polynucléotide de a) ou b) ; et codant une enzyme ayant une activité estérase ; et
e) un polynucléotide complémentaire de l'un quelconque des polynucléotides de a) à d).

2. Polynucléotide qui hybride dans des conditions stringentes à :

- a) un polynucléotide comprenant une séquence d'acide nucléique représentée dans SEQ ID NO : 23 ; ou
b) un polynucléotide complémentaire d'un polynucléotide de a) ;

et codant une enzyme ayant une activité estérase.

3. Polynucléotide selon la revendication 1 ou 2, dans lequel le polynucléotide est de l'ADN ou de l'ARN.

4. Vecteur comprenant le polynucléotide selon l'une quelconque des revendications 1 à 3.

5. Cellule hôte comprenant le vecteur selon la revendication 4.

6. Procédé de production d'un polypeptide comprenant l'expression, à partir de la cellule hôte selon la revendication 5, d'un polypeptide codé par ledit polynucléotide.

7. Procédé de production d'une cellule comprenant la transformation ou la transfection de la cellule avec le vecteur selon la revendication 4 de sorte que la cellule exprime le polypeptide codé par le polynucléotide contenu dans le vecteur.

8. Enzyme comprenant un membre choisi dans le groupe constitué de :

- a) une enzyme codée par le polynucléotide selon l'une quelconque des revendications 1(a), (b) ou (d) ;
b) une enzyme comprenant une séquence d'acides aminés identique à au moins 70%, 90% ou 95% à une séquence d'acides aminés représentée dans SEQ ID NO : 33 ;
c) une enzyme comprenant au moins 30 acides aminés de l'enzyme de (a) ou (b) ; et
d) une enzyme susceptible d'être obtenue par le procédé de la revendication 6.

9. Anticorps reconnaissant spécifiquement l'enzyme selon la revendication 8.

10. Méthode pour transférer un groupe amino d'un acide aminé à un α -céto acide, comprenant la mise en contact d'un acide aminé en présence d'un α -céto acide avec l'enzyme selon la revendication 8.

11. Méthode pour isoler l'enzyme selon la revendication 8, comprenant l'utilisation de l'anticorps selon la revendication 9.

12. Utilisation de l'enzyme selon la revendication 8 :

- a) comme inducteur de la maturation du fromage dans l'industrie laitière ;
b) pour éliminer la lignine dans l'industrie de la pâte et du papier ;
c) pour la synthèse de dérivés d'hydrates de carbone ; et/ou
d) pour la conversion de déchets ligno-cellulosiques en sucres fermentables.

13. Utilisation du polynucléotide de la revendication 1 ou 2 pour l'identification de séquences similaires dans d'autres organismes.

FIGURE 1

Staphylococcus marinus - 81-12LC

ATG TCT TTA AAC AAG CAC TCT TGG ATG GAT ATG ATA ATA TTT ATT CTC AGC TTT TCT TTC	60
Met Ser Leu Asn Lys His Ser Trp Met Asp Met Ile Ile Phe Ile Leu Ser Phe Ser Phe	20
CCA TTA ACA ATG ATC GCA TTA GCT ATC TCT ATG TCG TCA TGG TTT AAT ATA TGG AAT AAT	120
Pro Leu Thr Met Ile Ala Leu Ala Ile Ser Met Ser Ser Trp Phe Asn Ile Trp Asn Asn	40
GCA TTA AGC GAT CTA GGA CAT GCT GTT AAA AGC AGT GTT GCT CCA ATA TTC AAT CTA CGT	180
Ala Leu Ser Asp Leu Gly His Ala Val Lys Ser Ser Val Ala Pro Ile Phe Asn Leu Gly	60
CTT GCA ATT GGT GGG ATA CTA ATT GTT ATA GTT GGT TTA AGA AAT CTT TAT TCG TGG ACT	240
Leu Ala Ile Gly Gly Ile Leu Ile Val Ile Val Gly Leu Arg Asn Leu Tyr Ser Trp Ser	80
AGA GTT AAA GGA TCT TTA ATC ATA TCC ATG GGT GTA TTT CTT AAC TTA ATA GGG GTT TTC	300
Arg Val Lys Gly Ser Leu Ile Ile Ser Met Gly Val Phe Leu Asn Leu Ile Gly Val Phe	100
GAC GAA GTC TAT GGT TGG ATA CAT TTC CTA GTC TCA GTC TTG TTT TTC TTA TCA ATA ATA	360
Asp Glu Val Tyr Gly Trp Ile His Phe Leu Val Ser Val Leu Phe Phe Leu Ser Ile Ile	120
GCA TAT TTC ATA GCT ATA TCA ATA CTT GAC AAA TCA TGG ATA GCT GTT CTA CTA ATA ATA	420
Ala Tyr Phe Ile Ala Ile Leu Asp Lys Ser Trp Ile Ala Val Leu Leu Ile Ile	140
GGT CAT ATT GCA ATG TGG TAT CTA CAC TTT CCT TCA GAG ATT CCG AGA GGT GCG GCT ATT	480
Gly His Ile Ala Met Trp Tyr Leu His Phe Ala Ser Glu Ile Pro Arg Gly Ala Ala Ile	160
CCC GAG TTA TTA GCG GTA TTC TCG TTT TTA CCA TTC TAT ATA AGA GAC TAT TTT AAA TCA	540
Pro Glu Leu Leu Ala Val Phe Ser Phe Leu Pro Phe Tyr Ile Arg Asp Tyr Phe Lys Ser	180
TAC ACT AAA CGA TAG	555
Tyr Thr Lys Arg	185

FIGURE 2

Pyrodictium - TAG11-17LC

ATG AAA CTC CTT GAG CCC ACA AAT ACC TCC TAC ACG CTG TTA CAG GAT TTA GCA TTG CAT	60
Met Lys Leu Leu Glu Pro Thr Asn Thr Ser Tyr Thr Leu Leu Gln Asp Leu Ala Leu His	20
TTT GCA TTT TAC TCG TTT CTG GCC GTG TAT ACG TGG TTA CCC GGT GTC CTA GTC CGG CGC	120
Phe Ala Phe Tyr Trp Phe Leu Ala Val Tyr Thr Trp Leu Pro Gly Val Leu Val Arg Gly	40
GTA GCT GTG GAC ACA GGG GTG GCT CGG GTG CCT GGG CTC GGC CGG CGC GGT AAG AGG CTG	180
Val Ala Val Asp Thr Gly Val Ala Arg Val Pro Gly Leu Gly Arg Arg Gly Lys Arg Leu	60
CTC CTG GCC GCT GTG GCT GTC TTG GCG CTT GTT GTG CCT GTC CGG GCT TAT GTG	240
Leu Leu Ala Ala Val Leu Ala Val Val Val Ser Val Val Pro Ala Tyr Val	80
GCG TAT AGT AGT CTG CAC CCG GAG AGC TGT CGG CCC GTC GCG GAG CGG CTC ACC TAC	300
Ala Tyr Ser Ser Leu His Pro Glu Ser Cys Arg Pro Val Ala Pro Glu Gly Leu Thr Tyr	100
AAA GAG TTC AGC GTG ACC GCG GAG GAT GGC TGG GTG GTC CGG GGC TGG GTG CTG GGC CCC	360
Lys Glu Phe Ser Val Thr Ala Glu Asp Gly Leu Val Val Arg Gly Trp Val Leu Gly Pro	120
GC GCT GGG GGC AAC CCG GTG TPC GTT TPG ATG CAC GGG TAT ACT GGG TGC CGC TCG GCG	420
Gly Ala Gly Gly Asn Pro Val Phe Val Leu Met His Gly Tyr Thr Gly Cys Arg Ser Ala	140
CCC TAC ATG CCT GTG CTG GCC CGG GAG CTC GTG GAG TGG GGG TAC CGG GTG GTG TTC	480
Pro Tyr Met Ala Val Leu Ala Arg Glu Leu Val Glu Trp Gly Tyr Pro Val Val Val Phe	160
GAC TTC CGG GGC CAC GGG GAG AGC GGG GGC TCG ACG ACG ATT GGG CCC CGG GAG GTG CTG	540
Asp Phe Arg Gly His Gly Glu Ser Gly Ser Thr Thr Ile Gly Pro Arg Glu Val Leu	180
GAT GCC CGG CCT GTG GTC GGC TAT GTC TCG GAG CGG TTC CCC GGC CGC CGG ATA ATA TTG	600
Asp Ala Arg Ala Val Val Gly Tyr Val Ser Glu Arg Phe Pro Gly Arg Arg Ile Ile Leu	200
GTG GGG TTC AGT ATG GGC GCT GTA GCG ATC GTG GAG GGT GCT GGG GAC CGG CGG GTC	660
Val Gly Phe Ser Met Gly Ala Val Ala Ile Val Glu Gly Ala Gly Asp Pro Arg Val	220
TAC CGG GTG GCT GAT AGC CCG TAC TAT AGG CTC CGG GAC GTC ATA CCC CGG TGG CTG	720
Tyr Ala Val Ala Ala Asp Ser Pro Tyr Tyr Arg Leu Arg Asp Val Ile Pro Arg Trp Leu	240
GAG TAC AAG ACG CCG CTG CCG GGC TGG GTG GGT GTG CTG GCC CGG TTC TAC CGG AGG CTG	780
Glu Tyr Lys Thr Pro Leu Pro Gly Trp Val Gly Val Leu Ala Gly Phe Tyr Gly Arg Leu	260
ATG GCG CGC GTC GAC CTC CGC TTC CGC CCC GCT GGG GTG GAG CGG CGC GTC GAT AAG CGG TTG	840
Met Ala Gly Val Asp Leu Gly Phe Gly Pro Ala Gly Val Glu Arg Val Asp Lys Pro Leu	280
CTG GTG GTG TAT GGG CCC CGG GAC CCG CTG GTG ACC CGG GAC GAG CGG AGC AGC CTG CGG	900
Leu Val Val Tyr Gly Pro Arg Asp Pro Leu Val Thr Arg Asp Glu Ala Arg Ser Leu Ala	300
TCC CGT AGC CGG TGT GGC CGT CTC GTC GAG GTT CCT GGG GCT GGC CAC GTG GAG GCC GTG	960
Ser Arg Ser Pro Cys Gly Arg Leu Val Glu Val Pro Gly Ala Gly His Val Glu Ala Val	320
GAT GTG CTC CGG CGG CGC CGC TAC GCA GAC ATG CTG ATA GAG CTG GCG CAC GAG GAG TGC	1020
Asp Val Leu Gly Pro Gly Arg Tyr Ala Asp Met Leu Ile Glu Leu Ala His Glu Glu Cys	340
CCT CGG GGG GCC GGT GGC TGA	1041
Pro Pro Gly Ala Gly Gly	346

FIGURE 3

Archaeoglobus Venificus SN P6-24LC

ATG CCA TAT GTT AGG AAT GGT GGT GTC AAT ATC TAT TAT GAA CTG GTG GAT GGA CCT GAG Met Pro Tyr Val Arg Asn Gly Gly Val Asn Ile Tyr Tyr Glu Leu Val Asp Gly Pro Glu	60 20
CCA CCA ATT GTC TTT GTT CAC GGA TGG ACA GCA AAT ATG AAT TTT TGG AAA GAG CAA AGA Pro Pro Ile Val Phe Val His Gly Trp Thr Ala Asn Met Asn Phe Trp Lys Glu Gln Arg	120 40
CGT TAT TTT GCA GGC AGG AAT ATG ATG TTG TTT GTC GAT AAC AGA GGT CAT GGC AGG TCC Arg Tyr Phe Ala Gly Arg Asn Met Met Leu Phe Val Asp Asn Arg Gly His Arg Ser	180 60
GAT AAG CCA CTT GGA TAC GAT TTC TAC AGA TTT GAG AAC TTC ATT TCA GAT TTA GAT GCG Asp Lys Pro Leu Gly Tyr Asp Phe Tyr Arg Phe Glu Asn Phe Ile Ser Asp Leu Asp Ala	240 80
GTT GTT AGG GAG ACT GGA GTG GAG AAA TTT GTT CTC GTC GGA CAT TCA TTC GGA ACA ATA ATG Val Val Arg Glu Thr Gly Val Glu Lys Phe Val Leu Val Gly His Ser Phe Gly Thr Met	300 100
ATC TCT ATG AAG TAC TGT TCG GAG TAT CGG AAT CGG GTT CTT GCT CTA ATC CTC ATA GGT Ile Ser Met Lys Tyr Cys Ser Glu Tyr Arg Asn Arg Val Leu Ala Leu Ile Ile Lys	360 120
GGT GCG AGC AGA ATA AAG CTT CTA CAC AGA ATT GGA TAT CCT TTA GCA AAG ATT CTT GCA Gly Gly Ser Arg Ile Lys Leu Leu His Arg Ile Gly Tyr Pro Leu Ala Lys Ile Leu Ala	420 140
TCC ATT GCA TAC AAG AAG TCT TCA AGA TTG GTC GCA GAT CTT TCC TTT GGC AAA AAT GCT Ser Ile Ala Tyr Lys Ser Ser Arg Leu Val Ala Asp Leu Ser Phe Gly Lys Asn Ala	480 160
GGT GAA CTT AAA GAG TGG GGA TGG AAA CAG GCA ATG GAT TAT ACA CCC TCC TAC GTG GCA Gly Glu Leu Lys Glu Trp Gly Trp Lys Gln Ala Met Asp Tyr Thr Pro Ser Tyr Val Ala	540 180
ATG TAC ACG TAC AGA ACT CTA ACG AAA GTG AAT CTT GAA AAT ATC TTG GAG AAA ATA GAC Met Tyr Thr Tyr Arg Thr Leu Thr Lys Val Asn Leu Glu Asn Ile Leu Lys Ile Asp	600 200
TGT CCA ACA CTG ATT ATC GTT GGA GAA GAG GAT GCA CTA TTG CCC GTT AGC AAA TCA GTT Cys Pro Thr Leu Ile Ile Val Gly Glu Glu Asp Ala Leu Leu Pro Val Ser Lys Ser Val	660 220
GAG CTG AGC AGG AGG ATA GAA AAC TCA AAG CTT GTG ATC ATC CCA AAC TCG GGG CAT TGC Glu Leu Ser Arg Arg Ile Glu Asn Ser Lys Leu Val Ile Ile Pro Asn Ser Gly His Cys	720 240
GTA ATG CTT GAG AGT CCA AGT GAG GTT AAT AGA GCA ATG GAC GAA TTC ATT TCT TCA GCA Val Met Leu Glu Ser Pro Ser Glu Val Asn Arg Ala Met Asp Glu Phe Ile Ser Ser Ala	780 260
CAC TTC TAA Gin Phe	789 263

FIGURE 4

Aquifex pyrophilus - 28LC

TTG AGA TTG AGG AAA TTT GAA GAG ATA AAC CTC GTT CTT TCG GGA GCT GCA AAG GGC Leu Arg Leu Arg Lys Phe Glu Glu Ile Asn Leu Val Ser Gly Gly Ala Ala Lys Gly	60 20
ATA GCC CAC ATA CGT GTT TTG AAA GCT ATA AAC GAG CTC CGT ATA AGG GTG AGG GCT TTA Ile Ala His Ile Gly Val Leu Lys Ala Ile Asn Glu Leu Gly Ile Arg Val Arg Ala Leu	120 40
AGC GGG GTG AGC GCC GGG GCA ATC GTT TCG GTC TTT TAT GCC TCA GGC TAC TCC CCT GAA Ser Gly Val Ser Ala Gly Ala Ile Val Ser Val Phe Tyr Ala Ser Gly Tyr Ser Pro Glu	180 60
GGG ATG TTC AGC CTT CTG AAG AGG GTA AAC TGG CTG AAG CTG TTT AAG TTC AAG CCA CCT Gly Met Phe Ser Leu Leu Lys Arg Val Asn Trp Leu Lys Leu Phe Lys Pro Pro	240 80
CTG AAG GGA TTG ATA GGG TGG GAG AAG GCT ATA AGA TTC CTT GAG GAA GTT CTC CCT TAC Leu Lys Gly Leu Ile Gly Trp Glu Lys Ala Ile Arg Phe Leu Glu Glu Val Leu Pro Tyr	300 100
AGG AGA ATA GAA AAA CTT GAG ATA CCG ACG TAT ATA TGC GCG ACQ GAT TTA TAC TCG GGA Arg Arg Ile Glu Lys Leu Glu Ile Pro Thr Tyr Ile Cys Ala Thr Asp Leu Tyr Ser Gly	360 120
AGG CCT CTA TAC CTC TCG GAA GGG AGT TTA ATC CCC GCA CTT CTC GGC AGC TGT GCA ATT Arg Ala Leu Tyr Leu Ser Glu Gly Ser Leu Ile Pro Ala Leu Leu Gly Ser Cys Ala Ile	420 140
CCC GGC ATA TTT GAA CCC GTT GAG TAT AAG ATAT TAC TTG CTC GTT GAC GGA GGT ATA GTT Pro Gly Ile Phe Glu Pro Val Glu Tyr Lys Asn Tyr Leu Val Asp Gly Gly Ile Val	480 160
AAC AAC CTT CCC GTT GAG CCC TTT CAG GAA AGC GGT ATT CCC ACC GTT TGC GTT GAT GTC Asn Asn Leu Pro Val Glu Pro Phe Gln Glu Ser Gly Ile Pro Thr Val Cys Val Asp Val	540 180
CTT CCC ATA GAG CCG GAA AAG GAT ATA AAG AAC ATT CTT CAC ATC CTT TTG AGG AGC TTC Leu Pro Ile Glu Pro Glu Lys Asp Ile Lys Asn Ile Leu His Ile Leu Leu Arg Ser Phe	600 200
TTT CTT GCG GTC CGC TCA AAC TCC GAA AAG AGA AAG GAG TTT TGT GAC CTC GTT ATA GTT Phe Leu Ala Val Arg Ser Asn Ser Glu Lys Arg Lys Glu Phe Cys Asp Leu Val Ile Val	660 220
CCT GAG CTT GAG GAG TTC ACA CCC CTT GAT GTT AGA AAA GCG GAC CAA ATA ATG GAG AGG Pro Glu Leu Glu Glu Phe Thr Pro Leu Asp Val Arg Lys Ala Asp Gln Ile Met Glu Arg	720 240
GGA TAC ATA AAG GCC TTA GAG GTA CTT TCT GAA TAG Gly Tyr Ile Lys Ala Leu Glu Val Leu Ser Glu	756 252

FIGURE 5

M11TL-29L.

ATG TTT AAT ATC AAT GTC TTT GTT AAT ATA TCT TGG CTG TAT TTT TCA GGG AIA GTT ATG Met Phe Asn Ile Asn Val Phe Val Asn Ile Ser Trp Leu Tyr Phe Ser Gly Ile Val Met	60 20
AAG ACT GTG GAA GAG TAT GCG CTA CTT GAA ACA GGC GTA AGA GTG TTT TAT CGG TGT GTA Lys Thr Val Glu Glu Tyr Ala Leu Leu Glu Thr Gly Val Arg Val Phe Tyr Arg Cys Val	120 40
ATC CCG GAG AAA GCT TTT AAC ACT TTG ATA ATA GGT TCA CAC GGA TTG CGG GCG CAC AGT Ile Pro Glu Lys Ala Phe Asn Thr Leu Ile Ile Gly Ser His Gly Leu Gly Ala His Ser	180 60
GGA ATC TAC ATT AGT GTT GCT GAA GAA TTT GCT AGG CAC GGA TTT GGA TTC TGC ATG CAC Gly Ile Tyr Ile Ser Val Ala Glu Glu Phe Ala Arg His Gly Phe Gly Phe Cys Met His	240 80
GAT CAA AGG GGA CAT CGG AGA ACG GCA AGC GAT AGA GAA AGA GGG TAT GTG GAG GGC TTT Asp Gln Arg Gly His Gly Arg Thr Ala Ser Asp Arg Glu Arg Gly Tyr Val Glu Gly Phe	300 100
CAC AAC TTC ATA GAG GAT ATG AAG GCC TTC TCC GAT TAT GCC AAG TGG CGC GTG GGA GGT His Asn Phe Ile Glu Asp Met Lys Ala Phe Ser Asp Tyr Ala Lys Trp Arg Val Gly Gly	360 120
GAC GAA ATA ATA TTG CTA GGA CAC AGT ATG GGC GGG CTG ATA GCG CTC TTA ACA GTT GCA Asp Glu Ile Ile Leu Leu Gly His Ser Met Gly Leu Ile Ala Leu Leu Thr Val Ala	420 140
ACT TAT AAA GAA ATC GCC AAG GGA GTT ATC GCG CTA GCC CCG GGC CTC CAA ATC CCC TTA Thr Tyr Lys Glu Ile Ala Lys Gly Val Ile Ala Leu Ala Pro Ala Leu Gln Ile Pro Leu	480 160
ACC CGG GCT AGA AGA CTT GTT CTA AGC CTC CGC TCA AGG CTT GCC CGG CAT TCT AAG ATC Thr Pro Ala Arg Arg Leu Val Leu Ser Leu Ala Ser Arg Leu Ala Pro His Ser Lys Ile	540 180
ACC TTA CAA AGG AGA TTG CCG CAG AAA CCA GAG GGT TTT CAA AGA GCA AAA GAT ATA GAA Thr Leu Gln Arg Arg Leu Pro Gln Lys Pro Glu Gly Phe Gln Arg Ala Lys Asp Ile Glu	600 200
TAC AGT CTG AGT GAA ATA TCA GTC AAG CTC GTG GAC GAA ATG ATT AAA GCA TCA TCT ATG Tyr Ser Leu Ser Glu Ile Ser Val Lys Leu Val Asp Glu Met Ile Lys Ala Ser Ser Met	660 220
TTC TGG ACC ATA GCA GGG GAA ATT AAT ACT CCC GTC CTG CTT ATT CAT GGG GAA AAA GAC Phe Trp Thr Ile Ala Gly Glu Ile Asn Thr Pro Val Leu Leu Ile His Gly Glu Lys Asp	720 240
AAT GTC ATA CCT CCG GAG GCG AGC AAA AAA GCC TAC CAA TTA ATA CCT TCA TTC CCT AAA Asn Val Ile Pro Pro Glu Ala Ser Lys Lys Ala Tyr Gln Leu Ile Pro Ser Phe Pro Lys	780 260
GAG TTG AAA ATA TAC CCC GAT CTT CGA CAC AAC TTG TTT TTT GAA CCA GGC GCG GTG AAA Glu Leu Lys Ile Tyr Pro Asp Leu Gly His Asn Leu Phe Phe Pro Gly Ala Val Lys	840 280
ATC GTC ACA GAC ATT GTC GAG TGG GTT AAG AAT CTA CCC AGG GAA AAT CCT TAA Ile Val Thr Asp Ile Val Glu Trp Val Lys Asn Leu Pro Arg Glu Asn Pro	894 298

FIGURE 6

Thermococcus CL-2-3OLC

ATG GAG GTT TAC AAG GCC AAA TTC GGC GAA GCA AAG CTC CGC TGG GTC GTC GTT CTG GTT CAT Met Glu Val Tyr Lys Ala Lys Phe Gly Glu Ala Lys Leu Gly Trp Val Val Leu Val His	60 20
GGC CTC GGC GAG CAC AGC GGA AGG TAT GGA AGA CTG ATT AAG GAA CTC AAC TAT GGC GGC Gly Leu Gly Glu His Ser Gly Arg Tyr Gly Arg Leu Ile Lys Glu Leu Asn Tyr Ala Gly	120 40
TTT CGA GTT TAC ACC TTC GAC TGG CCC GGC CAC GGG AAG AGC CCG GGC AAG AGA CGG CAC Phe Gly Val Tyr Thr Phe Asp Trp Pro Gly His Gly Ser Pro Gly Lys Arg Gly His	180 60
ACG AGC GTC GAG GAG ATG GAA ATC ATC GAC TCG ATA ATC GAG GAG ATC AGG GAG AAG Thr Ser Val Glu Ala Met Glu Ile Ile Asp Ser Ile Ile Glu Glu Ile Arg Glu Lys	240 80
CCC TTC CTC TTC GGC CAC AGC CTC GGT GGT CTA ACT GTC ATC AGG TAC GCT GAG ACG CGG Pro Phe Leu Phe Gly His Ser Leu Gly Gly Leu Thr Val Ile Arg Tyr Ala Glu Thr Arg	300 100
CCC GAT AAA ATA CGG GGA TTA ATA GCT TCC TCG CCT GCC CTC GCC AAG AGC CCG GAA ACC Pro Asp Lys Ile Arg Gly Leu Ile Ala Ser Ser Pro Ala Leu Ala Lys Ser Pro Glu Thr	360 120
CCG GGC TTC ATG GTG GCC CTC GCG AAG TTC CTT GGA AAG ATC GCC CCG GGA GTT GTT CTC Pro Gly Phe Met Val Ala Leu Ala Lys Phe Leu Gly Lys Ile Ala Pro Gly Val Val Leu	420 140
TCC AAC GGC ATA ARG CCG GAA CTC CTC TCG AGG AAC AGC GAC GCC GTG AGG AGG TAC GTT Ser Asn Gly Ile Lys Pro Glu Leu Leu Ser Arg Asn Arg Asp Ala Val Arg Arg Tyr Val	480 160
GAA GAC CCA CTC GTC CAC GAC AGG ATT TCG GCC AAG CTC GGA AGG AGC ATC TTC GTG AAC Glu Asp Pro Leu Val His Asp Arg Ile Ser Ala Lys Leu Gly Arg Ser Ile Phe Val Asn	540 180
ATG GAG CTG GCC CAC AGG GAG GCG GAC AAG ATA AAA GTC CCG ATC CTC CTT CTG ATC GGC Met Glu Leu Ala His Arg Glu Ala Asp Lys Ile Lys Val Pro Ile Leu Leu Ile Gly	600 200
ACT GGC GAT GTA ATA ACC CCG CCT GAA GGC TCA CGC AGA CTC TTC GAG GAG CTG GCC GTC Thr Gly Asp Val Ile Thr Pro Pro Glu Gly Ser Arg Arg Leu Phe Glu Leu Ala Val	660 220
GAG AAC AAA ACC CTG AGG GAG TTC GAG GGG GCG TAC CAC GAG ATA TTT GAA GAC CCC GAG Glu Asn Lys Thr Leu Arg Glu Phe Glu Gly Ala Tyr His Glu Ile Phe Glu Asp Pro Glu	720 240
TGG GCC GAG GAG TTC CAC GAA ACA ATT GTT AAG TGG CTG GTT GAA AAA TCG TAC TCT TCG Trp Ala Glu Glu Phe His Glu Thr Ile Val Lys Trp Leu Val Glu Lys Ser Tyr Ser Ser	780 260
GCT CAA TAA Ala Gln	789 263

FIGURE 7

AquilaFax VP5-34LC

TTC ATT GGC AAT TTG AAA TTG AAG AGG TTT GAA GAG GTT AAC TTA GTT CTT TCG GGA GGG Leu Ile Gly Asn Leu Lys Leu Lys Arg Phe Glu Glu Val Asn Leu Val Leu Ser Gly Gly	60 20
GCT GCC AAG GGT ATC GCC CAT ATA GGT GTT TTA AAA GCT CTG GAA GAG CTC GGT ATA AAG Ala Ala Lys Gly Ile Ala His Ile Gly Val Leu Lys Ala Leu Glu Leu Gly Ile Lys	120 40
GTA AAG AGG CTC AGC GGG GTA AGT GCT GGA GCT ATC GTT TCC GTC TTT TAC GCT TCG GGC Val Lys Arg Leu Ser Gly Val Ser Ala Gly Ala Ile Val Ser Val Phe Tyr Ala Ser Gly	180 60
TAC ACT CCC GAC GAG ATG TTA AAA CTC CTG AAA GAG GTA AAC TGG CTC AAA CTT TTT AAG Tyr Thr Pro Asp Glu Met Leu Lys Leu Leu Lys Glu Val Asn Trp Leu Lys Leu Phe Lys	240 80
TTC AAA ACA CCG AAA ATG GGC TTA ATG GGG TGG GAG AAG GCT GCA GAG TTT TTG GAA AAA Phe Lys Thr Pro Lys Met Gly Leu Met Gly Trp Glu Lys Ala Ala Glu Phe Leu Glu Lys	300 100
GAG CTC GGA GTT AAG AGG CTG GAA GAC CTG AAC ATA CCA ACC TAT CTT TGC TCG GCG GAT Glu Leu Gly Val Lys Arg Leu Glu Asp Leu Asn Ile Pro Thr Tyr Leu Cys Ser Ala Asp	360 120
CTG TAC ACG GGA AAG GCT CTT TAC TTC GGC AGA GGT GAC TTA ATT CCC GTG CTT CTC GGA Leu Tyr Thr Gly Lys Ala Leu Tyr Phe Gly Arg Asp Leu Ile Pro Val Leu Leu Gly	420 140
AGT TGT TCC ATA CCC GGG ATT TTT GAA CCA GTT GAG TAC GAG AAT TTT CTA CTT GTT GAC Ser Cys Ser Ile Pro Gly Ile Phe Glu Pro Val Glu Tyr Glu Asn Phe Leu Leu Val Asp	480 160
GGA GGT ATA GTG AAC AAC CTG CCC GTA GAA CCT TTG GAA AAG TTC AAA GAA CCC ATA ATC Gly Ile Val Asn Asn Leu Pro Val Glu Pro Leu Lys Phe Lys Glu Pro Ile Ile	540 180
GGG GTA GAT GTG CTT CCC ATA ACT CAA GAA AGA AAG ATT AAA AAT ATA CTC CAC ATC CTT Gly Val Asp Val Leu Pro Ile Thr Glu Glu Arg Lys Ile Lys Asn Ile Leu His Ile Leu	600 200
ATA AGG AGC TTC TTT CTG GCG GTT CTC CCT AAT TCG GAA AAG AGA AAG GAG TTC TGC AAC Ile Arg Ser Phe Phe Leu Ala Val Arg Ser Asn Ser Glu Lys Arg Lys Glu Phe Cys Asn	660 220
GTA GTT ATA GAA CCT CCC CTT GAA GAG TTC TCT CCT CTG GAC GTA AAT AAG GCG GAC GAG Val Val Ile Glu Pro Pro Leu Glu Glu Phe Ser Pro Leu Asp Val Asn Lys Ala Asp Glu	720 240
ATA TTC TGC GGG GAT ATG AGA GCA CTT TAA Ile Phe Cys Gly Asp Met Arg Ala Leu	750 250

FIGURE 8

Teredinibacter - 42L

ATG CCA GCT AAT GAC TCA CCC ACG ATC GAC TTT AAT CCT CGC GGC ATT CTT CGC AAC GCT Met Pro Ala Asn Asp Ser Pro Thr Ile Asp Phe Asn Pro Arg Gly Ile Leu Arg Asn Ala	60 20
CAC GCA CAG GTT ATT TTA GCG ACT TCC GGC TTG CGC AAA GCG TTT TTG AAA CGC ACC CAC His Ala Gln Val Ile Leu Ala Thr Ser Gly Leu Arg Lys Ala Phe Leu Lys Arg Thr His	120 40
AAG AGC TAC CTC AGC ACT GCC CAA TGG CTG GAG CTC GAT GCC GGC AAC GGA GTT ACC TTG Lys Ser Tyr Leu Ser Thr Ala Gln Trp Leu Glu Leu Asp Ala Gly Asn Gly Val Thr Leu	180 60
GCC GGA GAG CCT AAC ACA GCG CCT GCA ACT GCA TCC TCC TCC CAC CGG GCG CAC AAG AAC Ala Gly Glu Leu Asn Thr Ala Pro Ala Thr Ala Ser Ser Ser His Pro Ala His Lys Asn	240 80
ACT CTG GTT ATT GTG CTG CAC GGC TGG GAA GGC TCC AGC CAG TCG GCC TAT GCG ACC TCC Thr Leu Val Ile Val Leu His Gly Trp Glu Gly Ser Ser Gln Ser Ala Tyr Ala Thr Ser	300 100
GCT GGC AGC ACG CCT TTC GAC AAT GGG TTC GAC ACT TTT CGC CCT AAT TTT CGC GAT CAC Ala Gly Ser Thr Leu Phe Asp Asn Gly Phe Asp Thr Phe Arg Leu Asn Phe Arg Asp His	360 120
GGC GAC ACC TAC CAC TTA AAC CGC GGC ATA TTT AAC TCA TCG CTG ATT GAC GAA GTA GTG Gly Asp Thr Tyr His Leu Asn Arg Gly Ile Phe Asn Ser Ser Leu Ile Asp Glu Val Val	420 140
GGC GCA GTC AAA GCC ATC CAG CAG CAA ACC GAC TAC GAC AAG TAT TGC CTG ATG GGG TTC Gly Ala Val Lys Ala Ile Gln Gln Thr Asp Tyr Asp Lys Tyr Cys Leu Met Gly Phe	480 160
TCA CTG GGT GGG AAC TTT GCC TTG CGC GTC GCG GCU GTG CGG GAA CAG CAT CTC GCT AAA CCG Ser Leu Gly Asn Phe Ala Leu Arg Val Ala Val Arg Glu Gln His Leu Ala Lys Pro	540 180
CTA CGG GGC GTG CTC GCC GTC TGC CCG GTC CTC GAC CCC GCA CAC ACC ATG ATG GCC CTA Leu Ala Gly Val Leu Ala Val Cys Pro Val Leu Asp Pro Ala His Thr Met Met Ala Leu	600 200
AAC CGA GGT GCG TTT TTC TAC GGC CGC TAT TTT GCG CAT AAA TGG AAG CGC TCG TTA ACC Asn Arg Gly Ala Phe Phe Tyr Gly Arg Tyr Phe Ala His Lys Trp Lys Arg Ser Leu Thr	660 220
GCA AAA CTT GCA GCT TTC CCA GAC TAC AAA TAC GGC AAA GAT TTA AAA TCG ATA CAC ACG Ala Lys Leu Ala Ala Phe Pro Asp Tyr Lys Tyr Gly Lys Asp Leu Lys Ser Ile His Thr	720 240
CTT GAT GAG TTA AAC AAC TAT TTC ATT CCC CGC TAC ACC GGC TTC AAC TCA GTC TCC GAA Leu Asp Glu Leu Asn Asn Tyr Phe Ile Pro Arg Tyr Thr Gly Phe Asn Ser Val Ser Glu	780 260
TAC TTC AAA AGT TAC ACG CTC ACC GGG CAG AAG CTC GCG TTT CTC AAC TGC CCC AGT TAC Tyr Phe Lys Ser Tyr Thr Leu Thr Gly Gln Lys Leu Ala Phe Leu Asn Cys Pro Ser Tyr	840 280
ATT CTG GCA GCT GGC GAC GAC CCA ATA ATT CCA GCA TCC GAC TTT CAG AAA ATA GCC AAG Ile Leu Ala Ala Gly Asp Asp Pro Ile Ile Pro Ala Ser Asp Phe Gln Lys Ile Ala Lys	900 300
CCT GCG AAT CTG CAC ATA ACA GTC ACG CAA CAA GGT TCT CAT TGC GCA TAC CTG GAA AAC Pro Ala Asn Leu His Ile Thr Val Thr Gln Gln Gly Ser His Cys Ala Tyr Leu Glu Asn	960 320
CTG CAF AAA CCT AGT GCT GGC GAC AAA TAT GCG GTG AAA TTA TTT GGA GCC TGT TGA Leu His Lys Pro Ser Ala Ala Asp Lys Tyr Ala Val Lys Leu Phe Gly Ala Cys End	1017 339

FIGURE 9

Archaeoglobus fulgidus VCl6 - 16MC1

ATG CTT GAT ATG CCA ATC GAC CCT GTT TAC TAC CAG CTT GCT GAG TAT
 Met Leu Asp Met Pro Ile Asp Pro Val Tyr Tyr Gln Leu Ala Glu Tyr
 1 5 10 15

TTC GAC AGT CTG CCG AAG TTC GAC CAG TTT TCC TCG GCC AGA GAG TAC
 Phe Asp Ser Leu Pro Lys Phe Asp Gln Phe Ser Ser Ala Arg Glu Tyr
 20 25 30

AGG GAG GCG ATA AAT CGA ATA TAC GAG GAG AGA AAC CGG CAG CTG AGC
 Arg Glu Ala Ile Asn Arg Ile Tyr Glu Glu Arg Asn Arg Gln Leu Ser
 35 40 45

CAG CAT GAG AGG GTT GAA AGA GTT GAG GAC AGG ACG ATT AAG GGG AGG
 Gln His Glu Arg Val Glu Arg Val Glu Asp Arg Thr Ile Lys Gly Arg
 50 55 60

AAC GGA GAC ATC AGA GTC AGA GTT TAC CAG CAG AAG CCC GAT TCC CCG
 Asn Gly Asp Ile Arg Val Arg Val Tyr Gln Gln Lys Pro Asp Ser Pro
 65 70 75 80

GGT CTG GTT TAC TAT CAC GGT GGT GGA TTT GTG ATT TGC AGC ATC GAG
 Val Leu Val Tyr His Gly Gly Phe Val Ile Cys Ser Ile Glu
 85 90 95

TCG CAC GAC GCC TTA TGC AGG AGA AYY GCG AGA CTT TCA AAC TCT ACC
 Ser His Asp Ala Leu Cys Arg Arg Ile Ala Arg Leu Ser Asn Ser Thr
 100 105 110

GTA GTC TCC GTG GAT TAC AGG CTC GCT CCT GAG CAC AAG TTT CCC CCC
 Val Val Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Ala
 115 120 125

CCA GTT TAT CAT TGC TAC GAT GCG ACC AAG TGG GTT GCT GAG AAC CGG
 Ala Val Tyr Asp Cys Tyr Asp Ala Thr Lys Trp Val Ala Glu Asn Ala
 130 135 140

GAG GAG CTG AGG ATT GAC CCG TCA AAA ATC TTC GTT GGG GGG GAC ACT
 Glu Glu Leu Arg Ile Asp Pro Ser Lys Ile Phe Val Gly Gly Asp Ser
 145 150 155 160

GCG GGA CGG AAT CTT GCC CCG GCG CTT TCA ATA ATG GCG AGA GAC AGC
 Ala Gly Asn Leu Ala Ala Ala Val Ser Ile Met Ala Arg Asp Ser
 165 170 175

GGA GAA GAT TTC ATA AAG CAT CAA ATT CTA ACT TAC CCC GTT GTG AAC
 Gly Glu Asp Phe Ile Lys His Gln Ile Leu Ile Tyr Pro Val Val Asn
 180 185 190

TTT GTA GCC CCC ACA CCA TCG CTT CTG GAG TTT GGA GAG GGG CTG TGG
 Phe Val Ala Pro Thr Pro Ser Leu Leu Glu Phe Gly Glu Gly Leu Trp
 195 200 205

ATT CTC GAC CAG AAG ATA ATG AGT TGG TTC TCG GAG CAG TAC TTC TCC
 Ile Leu Asp Gln Lys Ile Met Ser Trp Phe Ser Glu Gln Tyr Phe Ser
 210 215 220

AGA GAG GAA GAT AAG TTC AAG CCC CTC GCC TCC GTA ATC TTT GCG GAC
 Arg Glu Glu Asn Lys Phe Asn Pro Leu Ala Ser Val Ile Phe Ala Asp
 235 240 245 250

CTT GAG AAC CTA CCT CCT GCG CTG ATC ATA ACC GCC GAA TAC GAC CCG
Leu Glu Asn Leu Pro Pro Ala Leu Ile Ile Thr Ala Glu Tyr Asp Pro
255 260 265

CTG AGA GAT GAA GGA GAA GTT TTC GGG CAG ATG CTG AGA AGA GCC GGT
Leu Arg Asp Glu Gly Glu Val Phe Gly Gln Met Leu Arg Arg Ala Gly
270 275 280

GTT GAG GCG AGC ATC GTC AGA TAC AGA GGC GTG CTT CAC CGA TTC ATC
Val Glu Ala Ser Ile Val Arg Tyr Arg Gly Val Leu His Gly Phe Ile
285 290 295

AAT TAC TAT CCC GTG CTG AAG GCT GCG AGG GAT GCG ATA AAC CAG ATT
Asn Tyr Tyr Pro Val Leu Lys Ala Ala Arg Asp Ala Ile Asn Gln Ile
300 305 310

GCC GCT CTT CTT GTG TTC GAC TAG
Ala Ala Leu Leu Val Phe Asp
315 320

FIGURE 10

Sulfolobus Solfataricus P1 - SLC1

ATG CCC CTA GAT CCT AGA ATT AAA AAG TTA CTA GAA TCA GCT CTT ACT
 Met Pro Leu Asp Pro Arg Ile Lys Lys Leu Leu Glu Ser Ala Leu Thr
 1 5 10 15

ATA CCA ATT GGT AAA GCC CCA GTA GAA GAG GTA AGA AAG ATA TTT AGG
 Ile Pro Ile Gly Lys Ala Pro Val Glu Glu Val Arg Lys Ile Phe Arg
 20 25 30

CAA TTA GCG TCG GCA GCT CCC AAA GTC GAA GTT GGA AAA GTA GAA GAT
 Gln Leu Ala Ser Ala Ala Pro Lys Val Glu Val Gly Lys Val Glu Asp
 35 40 45

ATA AAA ATA CCA GGC AGT GAA ACC GTT ATA AAC GCT AGA GTG TAT TTT
 Ile Lys Ile Pro Gly Ser Glu Thr Val Ile Asn Ala Arg Val Tyr Phe
 50 55 60

CCG AAG AGT AGC GGT CCT TAT GGT GTT CTA GTG TAT CTT CAT GGA GGC
 Pro Lys Ser Ser Gly Pro Tyr Gly Val Leu Val Tyr Leu His Gly Gly
 65 70 75 80

GGT TTT GTA ATA GGC GAT GTG GAA TCT TAT GAC CCA TTA TGT AGA GCA
 Gly Phe Val Ile Gly Asp Val Glu Ser Tyr Asp Pro Leu Cys Arg Ala
 85 90 95

ATT ACA AAT GCG TGC AAT TGC GTT GTA GTA TCA GTG GAC TAT AGG TTA
 Ile Thr Asn Ala Cys Asn Cys Val Val Ser Val Asp Tyr Arg Leu
 100 105 110

GCT CCA GAA TAC AAG TTT CCT TCT GCA GTT ATC GAT TCA TTT GAC GCT
 Ala Pro Glu Tyr Lys Phe Pro Ser Ala Val Ile Asp Ser Phe Asp Ala
 115 120 125

ACT AAT TGG GTT TAT AAC AAT TTA GAT AAA TTT GAT GGA AAG ATG GGA
 Thr Asn Trp Val Tyr Asn Asn Leu Asp Lys Phe Asp Gly Lys Met Gly
 130 135 140

GTT GCG ATT GCG GGA GAT AGT GCT GGA GGA AAT TTG GCA GCG GTT GTA
 Val Ala Ile Ala Gly Asp Ser Ala Gly Gly Asn Leu Ala Val Val
 145 150 155 160

CCT CTT CTT TCA AAG GGT AAA ATT AAT TTG AAG TAT CAA ATA CTG GTT
 Ala Leu Leu Ser Lys Gly Lys Ile Asn Leu Lys Tyr Gln Ile Leu Val
 165 170 175

TAC CCA GCG GTA AGT TTA GAT AAC GTT TCA AGA TCC ATG ATA GAG TAC
 Tyr Pro Ala Val Ser Leu Asp Asn Val Ser Arg Ser Met Ile Glu Tyr
 180 185 190

TCT GAT GGG TTC TTC CTT ACC AGA GAG CAT ATA GAG TGG TTC GGT TCT
 Ser Asp Gly Phe Phe Leu Thr Arg Glu His Ile Glu Trp Phe Gly Ser
 195 200 205

CAA TAC TTA CGA AGC CCT GCA GAT TTG CTA GAC TTT AGG TTC TCT CCA
 Gln Tyr Leu Arg Ser Pro Ala Asp Leu Leu Asp Phe Arg Phe Ser Pro
 210 215 220

ATT CTG GCG CAA GAT TTC AAC GGA TTA CCT CCA GCC TTG ATA ATA ACA
Ile Leu Ala Gln Asp Phe Asn Gly Leu Pro Pro Ala Leu Ile Ile Thr
225 230 235 240

GCA GAA TAC GAT CCA CTA AGG GAT CAA GGA GAA GCG TAT GCA AAT AAA
Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Ala Tyr Ala Asn Lys
245 250 255

CTA CTA CAA GCT GGA GTC TCA GTT ACT AGT GTG AGA TTT AAC AAC GTT
Leu Leu Gln Ala Gly Val Ser Val Thr Ser Val Arg Phe Asn Asn Val
260 265 270

ATA CAC GGA TTC CTC TCA TTC TTT CCG TTG ATG GAG CAA GGA AGA GAT
Ile His Gly Phe Leu Ser Phe Pro Leu Met Glu Gln Gly Arg Asp
275 280 285

GCT ATA GGT CTG ATA GGG TCT GTG TTA AGA CGA GTA TTT TAT GAT AAA
Ala Ile Gly Leu Ile Gly Ser Val Leu Arg Arg Val Phe Tyr Asp Lys
290 295 300

ATT TAA
Ile
305

Figure 11
LAI1.1 Esterase es23

ATG	ARG	GTT	AAA	CAC	GTT	ATT	GTT	TTA	CAT	GTC	TTA	TAT	ATG	TCT	GCG	TTG	GTG	ATG	CGC	60
Met	Lys	Val	Lys	His	Val	Ile	Val	Leu	His	Gly	Leu	Tyr	Met	Ser	Gly	Leu	Val	Met	Arg	20
CGG	TTA	TGT	TGG	CGT	CTA	GAA	GAG	TGG	GGG	GTT	AAA	GTT	TTA	AAC	TTA	ACC	TAC	AAT	ACT	120
Pro	Leu	Cys	Ser	Arg	Leu	Glu	Glu	Ser	Gly	Val	Lys	Val	Leu	Asn	Leu	Thr	Tyr	Asn	Thr	40
CGA	GAC	CCT	AAT	CGA	GAT	GCT	ATT	TTT	ACG	CAA	ATA	GAT	GAG	TTT	ATT	AGC	AAT	GAG	CCT	180
Arg	Asp	Pro	Asn	Arg	Asp	Ala	Ile	Phe	Thr	Gln	Ile	Asp	Glu	Phe	Ile	Ser	Asn	Glu	Pro	60
TCT	GCT	TTA	GTG	TGT	CAC	TCT	ATG	GGG	GCG	TTA	GTT	GCT	CGC	TAT	TTA	GAG	GCA	AAC	Ser	240
Ser	Ala	Leu	Val	Cys	His	Ser	Met	Gly	Gly	Leu	Val	Ala	Arg	Ala	Tyr	Leu	Glu	Ala	Asn	80
TCA	GCG	CCA	AGT	CAT	CAT	CTT	GAA	AAG	GTA	ATC	ACC	TTA	GGA	ACG	CCA	CAT	ACT	GGC	AGC	300
Ser	Ala	Pro	Ser	His	His	Val	Glu	Lys	Val	Ile	Thr	Leu	Gly	Thr	Pro	His	Thr	Gly	Ser	100
CAT	ATT	GCT	GAA	AAA	ATG	CAG	CAA	AAA	GGG	TTC	GAG	CTA	TTA	AAA	AAT	AGC	GTT	GAG	360	
His	Ile	Ala	Glu	Lys	Met	Gln	Gln	Lys	Gly	Phe	Glu	Leu	Leu	Lys	Asn	Ser	Val	Glu	120	
TTT	TTA	CTC	TCT	AAG	AAT	GGT	GAT	TGG	CCT	TTT	AAA	GCT	AAG	CTA	TAT	AGC	ATT	GCC	GCG	420
Phe	Leu	Leu	Ser	Lys	Asn	Gly	Asp	Trp	Pro	Phe	Lys	Ala	Lys	Leu	Tyr	Ser	Ile	Ala	Gly	140
GAC	TTA	CCG	ATT	GGC	TTA	ATG	CCA	CTC	ATT	GTA	AAA	GGC	AGC	CGC	TCT	GAT	GGC	ACT	GTA	480
Asp	Leu	Pro	Ile	Gly	Leu	Met	Pro	Leu	Ile	Val	Lys	Gly	Ser	Arg	Ser	Asp	Gly	Thr	Val	160
TTG	CTA	GAT	GAA	ACC	AAG	CTA	AAG	GGT	ATG	GCT	GAA	CAC	AAG	GTG	TTT	CAT	TTA	AGC	CAT	540
Leu	Leu	Asp	Glu	Thr	Lys	Leu	Lys	Gly	Met	Ala	Glu	His	Lys	Val	Phe	His	Leu	Ser	His	180
ACA	AGT	ATG	ATT	TAC	TCT	CGC	CAA	GTC	GTT	AAT	TAT	ATT	CTT	GAG	CGC	TTG	AAC	GAG	GAC	600
Thr	Ser	Met	Ile	Tyr	Ser	Arg	Gln	Val	Val	Asn	Tyr	Ile	Leu	Glu	Arg	Leu	Asn	Glu	Asp	200
ATT	TA																		605	
Ile																			201	

Figure 12

Whale Mat Sample 11.801 Esterase e89

ATG ATA AAA AAC TTC GAC AGA GAA AAT TCT AGC TTA GTC CTG TCC GGT GGT GGT GCT CTC	60
Met Ile Lys Asn Phe Asp Arg Glu Asn Ser Ser Leu Val Leu Ser Gly Gly Ala Leu	20
GGT ATT GCT CAC TTG GGT GTA CTG CAT GAC CTT GAA AAA CAA AAT ATT GTA CCA AAT GAA	120
Gly Ile Ala His Leu Gly Val Leu His Asp Leu Glu Lys Gln Asn Ile Val Pro Asn Glu	40
ATT GTT GGT ACA AGT ATG GGT GGT ATC ATT GGT GCA TCT ATG GCT ATC GGG ATG AAA GAG	180
Ile Val Gly Thr Ser Met Gly Ile Ile Gly Ala Ser Met Ala Ile Gly Met Lys Glu	60
AAA GAA ATA CTC GAA GAA ATC AAA AAC TTT TCC AAT GTC TTC AAC TGG ATA AAA TTC TCT	240
Lys Glu Ile Leu Glu Ile Lys Asn Phe Ser Asn Val Phe Asn Trp Ile Lys Phe Ser	80
TTT TCC GGT AAT TCT GTT GTC GAT AAC GAG AAG ATC GCT AAG ATA TTT GAT ACT CTT TTT	300
Phe Ser Gly Asn Ser Val Val Asp Asn Glu Lys Ile Ala Lys Ile Phe Asp Thr Leu Phe	100
AAA GAC AGA AAG ATG ACA GAT ACC GTG ATC CCT CTT AAA CTC ATC GCT ACA AAC TTA CAT	360
Lys Asp Arg Lys Met Thr Asp Thr Val Ile Pro Leu Lys Leu Ile Ala Thr Asn Leu His	120
AAT GGA CAT AAA AAA GTC TTT ACT GCT TCG GAT GAT GTA CTG ATC AAA GAT GCA ATA CTC	420
Asn Gly His Lys Lys Val Phe Thr Ala Ser Asp Asp Val Leu Ile Lys Asp Ala Ile Leu	140
TCA ACA ATG GCA ATA CCC GGT GTA TTT GAA GAA CAT ATT ATT GAT GGT GAA ACC TAT GGC	480
Ser Thr Met Ala Ile Pro Gly Val Phe Glu Glu His Ile Ile Asp Gly Glu Thr Tyr Gly	160
GAC GGT TTT CTT TGT GAA AAC CTT GGT GTG AAT GAG GCA ACA TTC AAT GAT GTT TTA GCT	540
Asp Gly Phe Leu Cys Glu Asn Leu Gly Val Asn Glu Ala Thr Phe Asn Asp Val Leu Ala	180
GTA GAT GTC ATG GGT GAG AAC TCT TTT GAA AAA GCA ATG CCG GAC AAC TTC TTT AAA ACA	600
Val Asp Val Met Gly Glu Asn Ser Phe Glu Lys Ala Met Pro Asp Asn Phe Phe Lys Thr	200
TCA AAT GTT TTA GAA ATG TTT GAA AAA TCA ATG CGA CTT TTT ATT TAC AAC CAG ACA CAG	660
Ser Asn Val Leu Glu Met Phe Glu Lys Ser Met Arg Leu Phe Ile Tyr Asn Gln Thr Gln	220
ACA CAT ATT AAA AAT GCA AAT AAA AAT ATT TAT CTT ATT GAA CCC GTT ACC AAA GAG TAT	720
Thr His Ile Lys Asn Ala Asn Lys Asn Ile Tyr Leu Ile Glu Pro Val Thr Lys Glu Tyr	240
AAA ACA TTT CAA TTT CAT AAA CAT AAA GAG ATA CGT GCT TTA GGC TTG GGT TTA CTG TG	779
Lys Thr Phe Gln Phe His Lys Lys Glu Ile Arg Ala Leu Gly Leu Leu	259

Figure 13

Metallosphaera Prunata Rna 12/2 Esterase 23mc1

ATG CCC CTA CAT CCA AAG GTA AAG AAA TTA CTT TCC CAG CTA CCT CCC CAA GAC TTC TCC Met Pro Leu His Pro Lys Val Lys Lys Leu Leu Ser Gln Leu Pro Pro Gln Asp Phe Ser	60 20
AGA AAC GTG CAG GAC CTG AGG AAG GCC TGG GAT TTA CCC TTC TCA GGG AGG AGG GAG ACC Arg Asn Val Gln Asp Leu Arg Lys Ala Trp Asp Leu Pro Phe Ser Gly Arg Arg Glu Thr	120 40
CTG AAG AGG GTT GAG GAC CTT GAG ATA CCC ACT AGG GAC GCA CGA ATC AGG GCC AGG GTC Leu Lys Arg Val Glu Asp Leu Glu Ile Pro Thr Arg Asp Ala Arg Ile Arg Ala Arg Val	180 60
TAC ACC CCC TCA AGT AAG GAA AAC TTA CCC GTC CTT GTT TAC TAT CAC GGC GGT GGC TTC Tyr Thr Pro Ser Ser Lys Glu Asn Leu Pro Val Leu Val Tyr Tyr His Gly Gly Phe	240 80
GTC TTC GGT AGC GTT GAC AGC TAC GAC GGC CTC GCA TCC CTT ATT GCC AAG GAA TCT GGC Val Phe Gly Ser Val Asp Ser Tyr Asp Gly Leu Ala Ser Leu Ile Ala Lys Glu Ser Gly	300 100
ATT GCG GTT ATC TCC GTG GAG TAT AGG CTC GGC CCT GAG CAC AAG TTC CCC ACC GCA GTC Ile Ala Val Ile Ser Val Glu Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Thr Ala Val	360 120
AAC GAC TCG TGG GAT GCG CTT CTC TGG ATC GCG GAG AAC GGA GGC AAG CTG GGG CTC GAC Asn Asp Ser Trp Asp Ala Leu Leu Trp Ile Ala Glu Asn Gly Gly Lys Leu Gly Leu Asp	420 140
ACC TCG AGA CTT GCC GTG GCT GGG GAT AGT GCT GGA GGA AAC CTG TCT GCC GTC GTG TCC Thr Ser Arg Leu Ala Val Ala Gly Asp Ser Ala Gly Gly Asn Leu Ser Ala Val Val Ser	480 160
CTC CTG GAC AGG GAC CAG CGT AAG GGA CTG GTT ACT TAT CAG GTC CTA ATC TAC CCA GCA Leu Leu Asp Arg Asp Gln Gly Lys Glu Val Ser Tyr Gln Val Leu Ile Tyr Pro Ala	540 180
GTG AAC ATG GTC GAT AAC TCC CCA TCC GTC AGG GAG TAC CCG GAG GGA TAC TTC CTC ACC Val Asn Met Val Asp Asn Ser Pro Ser Val Arg Glu Tyr Gly Glu Gly Tyr Phe Leu Thr	600 200
AGG TCC ATG ATG AAC TCG TTC GGG ACC ATG TAC TTC TCC TCT GGA AGG GAA GCG GTC TCC Arg Ser Met Met Asn Trp Phe Gly Thr Met Tyr Phe Ser Gly Arg Glu Ala Val Ser	660 220
CCC TAC GGC TCT CCA GCC TTG GCT GAC CTA CAT AAC CTC CCA CCC TCA CTG GTG ATC ACT Pro Tyr Ala Ser Pro Ala Leu Ala Asp Leu Asn Leu Pro Pro Ser Leu Val Ile Thr	720 240
GCA GAG TAT GAT CCC CTA AGG GAT CAG GGA GAG ACC TAC TCT CAC TCC CTA AAC GAG GCT Ala Glu Tyr Asp Pro Leu Arg Asp Gln Gly Glu Thr Tyr Ser His Ser Leu Asn Glu Ala	780 260
GGA AAC GTA TCA ACC TTG GTT AGA TAT CAA GGA ATG ATT CAC GGC TTC CTG TCC TTC TAC Gly Asn Val Ser Thr Leu Val Arg Tyr Gln Gly Met Ile His Glu Phe Leu Ser Phe Tyr	840 280
GAG TGG ATA ACT GCC GGT AAA CTA GCC ATT CAC CAC ATT GCT GGG GTT CTG AGA TCT GTC Glu Trp Ile Thr Ala Gly Lys Leu Ala Ile His His Ile Ala Gly Val Leu Arg Ser Val	900 300
CTT TA Leu	905 301

Figure 14

Thermotoga neapolitana 5068 Esterase 56ec4

GTG GCC TTC TTC GAT ATG CCC CTT GAG GAA CTG AAA AAG TAC CGG CCT GAA AGG TAC GAG	60
Val Ala Phe Phe Asp Met Pro Leu Glu Leu Lys Tyr Arg Pro Glu Arg Tyr Glu	20
GAG AAA GAT TTC GAT GAG TTC TGG AGG GAA ACA CTT AAA GAA AGC GAA GCA TTC CCT CTG	120
Glu Lys Asp Phe Asp Glu Phe Trp Arg Glu Thr Leu Lys Ser Glu Gly Phe Pro Leu	40
GAT CCC GTC TTT GAA AAG GTG GAC TTT CAT CTC AAA ACC GTT GAA AGC TAC GAT GTT ACT	180
Asp Pro Val Phe Glu Lys Val Asp Phe His Leu Lys Thr Val Glu Thr Tyr Asp Val Thr	60
TTC TCT GGA TAC AGG GCG CAG AGA ATA AAG GGC TGG CTT CTT GTT CCC AMG TTG GCG GAA	240
Phe Ser Gly Tyr Arg Gly Gln Arg Ile Lys Gly Trp Leu Leu Val Pro Lys Leu Ala Glu	80
GAA AAG CTT CCA TGC GTG CAG TAC ATA GGT TAC AAT GGT GGA AGG GGT TTT CCA CAC	300
Glu Lys Leu Pro Cys Val Val Gln Tyr Ile Gly Tyr Asn Gly Gly Arg Gly Phe Pro His	100
GAC TGG CTG TTC TGG CCG TCA ATG GGT TAC ATC TGT TTT GTC ATG GAC ACC AGG GGG CAG	360
Asp Trp Leu Phe Trp Pro Ser Met Gly Tyr Ile Cys Phe Val Met Asp Thr Arg Gly Gln	120
GGA AGC GGC TGG ATG AAG GGA GAC ACA CCG GAT TAC CCT GAG GGT CCA GTC GAT CCA CAG	420
Gly Ser Gly Trp Met Lys Gly Asp Thr Pro Asp Tyr Pro Glu Gly Pro Val Asp Pro Gln	140
TAC CCC GGA TTC ATG ACG AGG GGC ATT CTG GAT CCC GGA ACC TAT TAC TAC AGG CGA GTC	480
Tyr Pro Gly Phe Met Thr Arg Gly Ile Leu Asp Pro Gly Thr Tyr Tyr Arg Arg Val	160
TTC GTG GAT GCG GTC AGG GCG GTG GAA GCA GCC ATT TCC TTC CCG AGA GTG GAT TCC AGG	540
Phe Val Asp Ala Val Arg Ala Val Glu Ala Ala Ile Ser Phe Pro Arg Val Asp Ser Arg	180
AAG GTG GTG GTG GCC GGA GGC AGT CAG GGT GGG GGA ATC CCC CTT GCG GTG AGT GCC CTG	600
Lys Val Val Val Ala Gly Ser Gln Gly Gly Ile Pro Leu Ala Val Ser Ala Leu	200
TCG AAC AGG GTG AAG GCT CTG CTC TGC GAT GTG CCG TTT CTG TGC CAC TTC AGA AGG GCC	660
Ser Asn Arg Val Lys Ala Leu Leu Cys Asp Val Pro Phe Leu Cys His Phe Arg Arg Ala	220
GTG CAA CTT GTC GAC ACA CAC CCA TAC GTG GAG ATC ACC AAC TTC CTC AAA ACC CAC AGG	720
Val Gln Leu Val Asp Thr His Pro Tyr Val Glu Ile Thr Asn Phe Leu Lys Thr His Arg	240
GAC AAA GAG GAG ATT GTT TTC AGA ACA CTT TCC TAC TTC GAT GGT GTG AAC TTT GCA GCA	780
Asp Lys Glu Glu Ile Val Phe Arg Thr Leu Ser Tyr Phe Asp Gly Val Asn Phe Ala Ala	260
AGG GCA AAG GTG CCC GCC CTG TTT TCC GTT GGG CTC ATG GAC ACC ATC TGT CCT CCC TCG	840
Arg Ala Lys Val Pro Ala Leu Phe Ser Val Gly Leu Met Asp Thr Ile Cys Pro Pro Ser	280
ACG GTC TTC GCC GCT TAC AAC CAC TAC GCC GGT CCA AAG GAG ATC AGA ATC TAT CCG TAC	900
Thr Val Phe Ala Ala Tyr Asn His Tyr Ala Gly Pro Lys Glu Ile Arg Ile Tyr Pro Tyr	300
AAC AAC CAC GAA GGT GGA GGT TCT TTC CAG GCA ATT GAG CAG GTG AAA TTC TTG AAG AGA	960
Asn Asn His Glu Gly Gly Ser Phe Gln Ala Ile Glu Gln Val Lys Phe Leu Lys Arg	320
CTA TTT GAG GAA GGC TAG	978
Leu Phe Glu Glu Gly	326

Figure 15

Malittangium lichenicola Esterase 77mcL

ATG CGC ACC CTC TCC TTC GGT CGG ATG ACC ACA GGG GGA AGC ATT CAC ATG GCG ACC ATG Met Arg Thr Leu Ser Phe Gly Pro Met Thr Thr Gly Gly Ser Ile His Met Ala Thr Met	60 20
GAC GTG ATG CGC GGG CGG CGG ATG CAG CGG CTG TCA CAG GGC GCC AGG GAG GCG GCG AAC Asp Val Met Arg Gly Pro Gly Met. Gln Arg Leu Ser Gln Gly Ala Arg Glu Ala Asn	120 40
CAC CCC TGG GCG AAG CGA CTG CGC CCG ATG GCG TAC GCG GCG AAC CGC GCG CTG TAC GCG His Pro Trp Ala Lys Arg Leu Gly Arg Met Gly Tyr Ala Ala Lys Gly Ala Val Tyr Ala	180 60
ATC ATC GGC GTG CTC GCG CTG AAG CTC GCG GCG GAG GGC GCG CGG ACC ACC AGC GAC ACC Ile Ile Gly Val Leu Ala Leu Lys Leu Ala Ala Gly Glu Gly Arg Thr Thr Asp Ser	240 80
CAC GCG GCG GTG AAC ACC GTG GCG CAC GCG CCC TTC GGC GTC GCG CTG CTU GCG GTG CTG His Gly Ala Val Asn Thr Val Ala His Gly Pro Phe Gly Val Ala Leu Leu Ala Val Leu	300 100
CTG CTG GGC CTG CTG GCG TAC GTG GTC TGG AGG TTC GGC CAG GCG CCC TTC GTC GAC ACC GAG Val Val Gly Leu Leu Gly Tyr Val Val Trp Arg Phe Ala Gln Ala Phe Val Asp Thr Glu	360 120
GAC AAG GGC TCC GAC GCG AAG GGA ATC GCG ACC CGC GCC ATG TAC TTC CTC AGC GGC TGC Asp Lys Gly Ser Asp Ala Lys Gly Ile Ala Thr Arg Ala Met Tyr Phe Leu Ser Gly Cys	420 140
ATC TAC GCG TCG CTG GCG CCC TTC GCG GCG CAG TCC CTG GTC GGC GCG GCG GCG CAC CGG Ile Tyr Ala Ser Leu Ala Phe Ala Ala Gln Ser Leu Val Gly Ala Ala His Gly Arg	480 160
AAC AAG GCG ACC CAG GCG TCG AGC GCG ACC CTC ARG GAG CAG CCC TTT GGC CGC CTG CTC Ser Lys Gly Thr Gln Gly Trp Thr Ala Thr Leu Met Glu Gln Pro Phe Gly Arg Val Leu	540 180
GTG GCG CTG GTG GGG CTG GGC ATC GTG GGC TTC CGG CTG AAG CAG TTC CAC ACC GCG TGG Val Ala Leu Val Gly Leu Gly Ile Val Gly Phe Ala Leu Lys Gln Phe His Thr Ala Trp	600 200
AAG GCG AAG TTC CGG GAG AAG CTC ACC CTC ACC GGA CTG CCT GCG CGG AAG CAG CAC CAC Lys Ala Lys Phe Arg Glu Lys Leu Thr Leu Thr Gly Leu Ala Ala Arg Lys Gln His His	660 220
ATC GAG CCC ATG TCC CAG TTC GGC ATC GCG CCC GCG GCG GTC GTC TTC GCG GTC ATC GGC Ile Glu Arg Met Cys Gln Phe Gly Ile Ala Ala Arg Gly Val Val Phe Ala Val Ile Gly	720 240
GCG TTC CTC GTC CGC TCC GCG CTG GAC GCG AAC CCC GCG GAG GCG GCG CTG GCG GAG Gly Phe Leu Val Arg Ser Ala Val Asp Ala Asn Pro Phe Gln Ala Lys Gly Leu Gly Glu	780 260
GCG CTG GCG GTC GTC GCG AGG CAG CGG TCC GGC GAC GTG CTC CTG CGG GTG GTC GCG GCG Ala Leu Ala Val Val Ala Arg Gln Pro Ser Gly Asp Val Leu Leu Gly Val Val Ala Ala	840 280
GCG CTG CTG GCG TAC GCG GCG TAC CTC TTC CTC CAG CGG CGC TAC CGC GAA CTC TAG Gly Leu Val Ala Tyr Ala Ala Tyr Leu Phe Leu Gln Ala Arg Tyr Arg Glu Leu End.	897 299

Figure 16

Whale Mat Sample 11.801 Esterase ee2

ATG ACC AAA TTC GCA ATA CTC TGG CGG TTG ATA ACG GCA TAC CTG CGG GAA CCT GTG ATG Met Ser Lys Phe Ala Ile Leu Trp Ala Leu Ile Thr Ala Tyr Leu Pro Glu Pro Val Met	60 20
AAA CTG GTA TAT TTA CGG CGG CGC GAA ACU CTT CGG GCA CGG ACU CTT GAC GTT AAA GCC Lys Leu Val Tyr Leu Gly Arg Arg Glu Thr Leu Gly Ala Arg Thr Leu Asp Val Lys Ala	120 90
CAA GCT GTC CGG CGG CGC GGC AAT GCA ACA AGA CCT GTC CGG GTG ATT CCG ACG GTC GAG Gln Ala Val Gly Arg Leu Ala Asn Ala Thr Arg Pro Val Gly Val Ile Pro Thr Val Glu	180 60
GAA AGC CGG AAG ATG ACG GAT AAA GCC GTT AGC CTT TTT GAT CAG CCC GCC CCC GAA TTA Glu Ser Arg Lys Met Thr Asp Lys Ala Val Ser Leu Phe Asp Gln Pro Ala Pro Glu Leu	240 80
TTC CGT AAA AAA GAC ATT CAG ATT GAC GGG GCT GAA CGG CCT ATT GAT GCC CGT ATT TAC Phe Arg Lys Lys Asp Ile Gln Ile Asp Gly Ala Glu Gly Pro Ile Asp Ala Arg Ile Tyr	300 100
AGC GGC CCT GCA AAA CAT CGC CCN CGN CCA ATN CTA GTG TAT TTT CAC GGC GGT GGC TGG Ser Gly Pro Ala Lys His Arg Pro Arg Pro Ile Leu Val Tyr Phe His Gly Gly Gly Trp	360 120
GTT CAG GGC AAT CTC GAC AGC CAT GAC GGG GTT TGC GGC AAG CTG GCA AAA TGG GCG AAC Val Gln Gly Asn Leu Asp Ser His Asp Gly Val Cys Gly Lys Leu Ala Lys Trp Ala Asn	420 140
TGC ATT GTT ATC TCG GTC GAT TAT CGT CTA CGG CCC GAA CAC AAA TTT CCT TGT GCG CCG Cys Ile Val Ile Ser Val Asp Tyr Arg Leu Ala Pro Glu His Lys Phe Pro Cys Ala Pro	480 160
CTT GAT GCG ATT CGG GCC TAT AAA TGG GTG CGC GCC AAC GCA ACA AAC CTT GGC GGC GAT Leu Asp Ala Ile Ala Ala Tyr Lys Trp Val Arg Ala Asn Ala Thr Asn Leu Gly Gly Asp	540 180
CCT GAA CGT ATC GGC GTT GGC GGC GAT AGC GCA GGG AAT CTT GCC GGC GTT GTC TGC Pro Glu Arg Ile Gly Val Gly Asp Ser Ala Gly Asn Leu Ala Ala Val Val Cys	600 200
CAA CAA ACC GCG ATG AAC GGC GAG CGC ACA CCA GAT CTG CAA GTC CTG ATC TAT CCG GCG Gln Gln Thr Ala Met Asn Gly Glu Arg Thr Pro Asp Leu Gln Val Leu Ile Tyr Pro Ala	660 220
CTG GAT GCA CGC ATG ATC TCG ACC TCG ATG GAG GAA TTG CGT GAT GCC TAC ATC TTG CGG Leu Asp Ala Arg Met Ile Ser Thr Ser Met Glu Leu Arg Asp Ala Tyr Ile Leu Pro	720 240
AAA TCC AGA ATG GAG TAT TTC CTC CGC CTA TAT AGC CGT GGC CCT GAC GAT ATC GAG GAC Lys Ser Arg Met Glu Tyr Phe Leu Gly Leu Tyr Thr Arg Gly Pro Asp Asp Ile Glu Asp	780- 260
CTT AGG ATG TCG CCA ATT CTC AGG GAT ACC GTC CGG GAT CAA CCC CAA GGC TGC ATT GTC Leu Arg Met Ser Pro Ile Leu Arg Asp Thr Val Ala Asp Gln Pro Gln Ala Cys Ile Val	840 280
ACC TGT GGG TTT GAC CCT GCG CGA CGA CGG GAA CAC CTA CGC CAA AGG CTT ATT TGC CGA Thr Cys Gly Phe Asp Pro Ala Arg Arg Glu His Leu Arg Arg Thr Leu Asn Cys Arg	900 300
GGG GAT AGA CGT TA Gly Asp Arg Arg	914 304

Figure 17

Whale Mat Sample AD3059 Esterase est4

GTG AGC ATT CGT CTG CGA CTG TTA AAC TGG TTT TTG AAT ACC TTT GAA AAA CCA AAA CTG	60
Val Ser Ile Arg Leu Arg Leu Leu Asn Trp Phe Leu Asn Thr Phe Glu Lys Pro Lys Leu	20
GCC GCG GCC AAA ACG CCG GAT GAT TTG CGA AAA TCG TTT GAA TTA AAG GCG AGG TTT TTG	120
Ala Ala Ala Lys Thr Pro Asp Asp Leu Arg Lys Ser Phe Glu Leu Lys Ala Arg Phe Leu	40
TTT CCG GCG CCA CGT AAA ACA AGG TTT AGT CAT GAT GTA TTG CAG TCA GGC ATC GGG TTG	180
Phe Pro Ala Pro Arg Lys Thr Arg Phe Ser His Asp Val Leu Gln Ser Gly Ile Gly Ser	60
GTA AAT GCC CAG TGG GCG AAA TCC AAA TCT GCA TCT GAT GAC AGG GTA ATC CTG TAT TTT	240
Val Asn Ala Gln Trp Ala Lys Ser Lys Ser Ala Ser Asp Asp Arg Val Ile Leu Tyr Phe	80
CAT GGG GGA GGG TAT GTT TTT GGG TCA CCA AAA ACG CAC CGT GCA ATG TTG GCG CCC TTG	300
His Gly Gly Tyr Val Phe Gly Ser Pro Lys Thr His Arg Ala Met Leu Ala Arg Leu	100
TGG GCA ATG ACA GGT CTT TCT GCG CTT CCA GAT TAT AGG TTG GCA CCA GAG CAC CCA	360
Ser Ala Met Thr Gly Leu Ser Ala Cys Leu Pro Asp Tyr Arg Leu Ala Pro Glu His Pro	120
TTT CCA GCC GCG ATC GAA GAT GCA GTT TTA TCG TAT AAA TGT TTA CTA GAG CGA GCA ATC	420
Phe Pro Ala Ala Ile Glu Asp Ala Val Leu Ser Tyr Lys Cys Leu Leu Glu Arg Ala Ile	140
GAG CCC CAA AAT ATT ATA CTG GGG GGG GAC ACT GCT GGT GGC GGT TTG GTT CTT GCT TTG	480
Glu Pro Gln Asn Ile Ile Leu Gly Gly Asp Ser Ala Gly Gly Leu Val Leu Ala Leu	160
CTT GCA GAA ATC AAG GCC CAA TCC TTG CCC AAA CCT GCT GGC GTT TTT GCC TTG TCG CCT	540
Leu Ala Glu Ile Lys Ala Gln Ser Leu Pro Lys Pro Ala Gly Val Phe Ala Leu Ser Pro	180
TTG GTT GAT TTA TCA TTT TCG GGC CTT TCG TTT TCT AAA AAT GCC CAA ACC GAT GTG ATG	600
Leu Val Asp Leu Ser Phe Ser Gly Leu Ser Phe Ser Lys Asn Ala Gln Thr Asp Val Met	200
TTG CCC GCA TCA CGG GCT GCG GAT ATG GCG ACC TTG TAT TTG GAT GCG GCC GAT GCA GAT	660
Leu Pro Ala Ser Arg Ala Ala Asp Met Ala Thr Leu Tyr Leu Asp Gly Ala Asp Ala Asp	220
GAT CCA CGT GCA TCG CGG CTC CAG GCG GAT TTT TCT GGC ATG CCG CCT GTC TTT CTG ACA	720
Asp Pro Arg Ala Ser Pro Leu Gln Ala Asp Phe Ser Gly Met Pro Pro Val Phe Leu Thr	240
GCA AGT GAC AGT GAA ATC CTG TTG GAT GAT TGC CTC CGG ATG GCG GAT CAC TTG CGT GCG	780
Ala Ser Asp Ser Glu Ile Leu Leu Asp Asp Cys Leu Arg Met Ala Asp His Leu Arg Ala	260
CAA GGT GTC GTT GTG ACA GAC CGG ATT GTT GAA AAC CAT CCA CAT GTC TTG CAT ATT TTT	840
Gln Gly Val Val Val Thr Asp Arg Ile Val Glu Asn His Pro His Val Trp His Ile Phe	280
CAA CGC CTT CTA CCC GAA GCA GAT CAG CGG CTC CGG CGG ATT GCC CGG TTG ATT AAA CCT	900
Gln Arg Leu Leu Pro Glu Ala Asp Gln Gly Leu Arg Ala Ile Ala Ala Trp Ile Lys Pro	300
CTT TTA TCA GGT TCA AAC GAA AGC TA	926
Leu Leu Ser Gly Ser Asn Glu Ser	308

Figure 18

Microscilla furvescens Esterase 53ec2

ATG CTT ACA TTT AAT GTT TTA TAT GGT ATG ATG AAA CAA AAA CTA GCA GCA ATT CTC ATG Met Leu Thr Phe Asn Val Leu Tyr Gly Met Met Lys Gln Lys Leu Ala Ala Ile Leu Met	60 20
TTT TTA GGG CTA TCA GCA GCA GAG GCT CAA GAC TGG CCT GAC CTA CAG AAA TAT CGT AGT Phe Leu Gly Leu Ser Ala Ala Gln Ala Gln Asp Trp Pro Asp Leu Gln Lys Tyr Arg Ser	120 40
GCT AAT AAA GAA GCC AAA TTA CTT CCA AAG GAA AAC CGG AAG GTG GTT TTT ATG GGC AAC Ala Asn Lys Glu Ala Lys Leu Leu Pro Lys Glu Asn Arg Lys Val Val Phe Met Gly Asn	180 60
TCC ATT ACA GAA GCC TGG ATT AGT CGG CGA CCT GAG TTT TTT ATG GAA AAT GGG TTT ATC Ser Ile Thr Glu Ala Trp Ile Ser Gln Arg Pro Glu Phe Phe Ser Glu Asn Gly Phe Ile	240 60
GGT CGA GGC ATC AGT GGC CGG ACA ACC CCT CAG ATG TTG TTG AGA TTC CGA CGG GAT GTG Gly Arg Gly Ile Ser Gly Gln Thr Thr Pro Gln Met Leu Leu Arg Phe Arg Gln Asp Val	300 100
ATA GAC CTG CAG CCA AAG GCT GTA GTG ATA CTA GCT GGT ACC AAT GAC GTA GCT CAA AAT Ile Asp Leu Gln Pro Lys Ala Val Val Ile Leu Ala Gly Thr Asn Asp Val Ala Gln Asn	360 120
ACC GGG CGG ATG ACC ATT GAG GAA TCG CTT GCT AAC ATT AAG TCT ATG GTG GAG CTG GCG Thr Gly Pro Met Thr Ile Glu Glu Ser Leu Ala Asn Ile Lys Ser Met Val Glu Leu Ala	420 140
CAA GCC AAT GGG ATC ACG CCT GTT TTG TGT ACC GTG CTG CCT GCA GAT CGT TTC AGC TGG Gln Ala Asn Gly Ile Thr Pro Val Leu Cys Thr Val Leu Pro Ala Asp Arg Phe Ser Trp	480 160
CGA CCT GAG CTT ACA CCC GCA GAA ACT ATC ATT GCC CTC AAT CAG CTC ATT AAG CAA TAT Arg Pro Glu Leu Thr Pro Ala Glu Thr Ile Ile Ala Leu Asn Gln Leu Ile Lys Gln Tyr	540 180
GCC GAG GCA CGG CCC CTG GCG GAT TAT CAT GCT GCA CTC ACC AAT AAA GGT GGA Ala Glu Ala Gln Gly Leu Ala Leu Val Asp Tyr His Ala Ala Leu Thr Asn Lys Gly Gly	600 200
GGA CTT CCU GTG AAA TAC GGA GAA GAT GGT GTG CAT CCA AAT GTA GCA GGC TAT CAG GTG Gly Leu Pro Val Lys Tyr Gly Glu Asp Gly Val His Pro Asn Val Ala Gly Tyr Gln Val	660 220
ATG GAA AAC ATT GTT TTA CCG GTC ATT TCC AGC GAG TTG GCA AAG CTG AAG TA Met Glu Asn Ile Val Leu Pro Val Ile Ser Ser Glu Leu Ala Lys Leu Lys	713 237

Figure 19

Thermotoga maritima MSB8 Esterase 6scl

ATG	GCC	TTC	TTC	GAT	TTA	CCA	CTC	GAA	GAA	CTG	AAG	AAA	TAT	CGT	CCA	GAG	CGG	TAC	GAA		60	
Met	Ala	Phe	Phe	Asp	Leu	Pro	Leu	Glu	Glu	Leu	Lys	Lys	Tyr	Arg	Pro	Glu	Arg	Tyr	Glu		20	
GAG	AAA	GAC	TTC	GAT	GAG	TTC	TGG	GAA	GAG	ACA	CTC	GCA	GAG	AGC	GAA	AAG	TTC	CCC	TTA		120	
																					40	
GAC	CCC	GTC	TTC	GAG	AGG	ATG	GAG	TCT	CAC	CTC	AAA	ACA	GTC	GAA	GCG	TAC	GAT	GTC	ACC		180	
																					60	
TTC	TCC	GGA	TAC	AGG	GGA	CAG	AGG	ATC	AAA	GGG	TGG	CTC	CTT	GTT	CCA	AAA	CTG	GAA	GAA		240	
																					80	
GAA	AAA	CTT	CCC	TGC	GTT	GTC	CAG	TAC	ATA	GGA	TAC	AAA	AAC	GGT	GGA	AGA	GGA	TTC	CCT	CAC		300
																					100	
GAC	TGG	CTG	TTC	TGG	CCT	TCT	ATG	GGT	TAC	ATA	TGT	TTC	GTC	ATG	GAT	ACT	CGA	GGT	CAG		360	
																					120	
GGA	AGC	GCC	TGG	CTG	AAA	GGA	GAC	ACA	CCG	GAT	TAC	CCT	GAG	GGT	CCC	GTT	GAC	CCT	CAG		420	
																					140	
TAT	CCA	GGA	TTC	ATG	ACA	AGA	GGA	ATA	CTG	GAT	CCC	AGA	ACT	TAC	TAC	TAC	AGA	CAG	TTC		480	
																					160	
TTC	ACG	GAC	GCT	GTC	AGA	GCC	GCC	GTT	GAA	GCT	GCT	GCT	TCT	TTT	CCT	CAG	GTA	GAT	CAG		540	
																					180	
AGA	ATC	GTG	ATA	GCT	GGA	GCG	AGT	CAG	GGT	GCC	GGA	ATA	GCC	CTT	GCG	GTG	AGC	GCT	CTC		600	
																					200	
TCA	AAG	AAA	GCA	AAG	GCT	CTT	CTG	TGC	GAT	GTG	CCG	TTT	CTG	TGT	CAC	TTC	AGA	AGA	GCA		660	
																					220	
GTA	CAG	CTT	GTG	GAT	ACG	CAT	CCA	TAC	CCG	GAG	ATC	ACG	AAC	TTT	CTA	AAG	ACC	CAC	AGA		720	
																					240	
GAC	AAG	GAA	GAA	ATC	GTG	TTC	AGG	ACT	CTT	TCC	TAT	TTC	GAT	GGA	GTG	AAC	TTC	GCA	GCC		780	
																					260	
AGA	GCG	AAG	ATC	CCT	GCG	CTG	TTT	TCT	GTG	GGT	CTC	ATG	GAC	AAC	ATT	TGT	CCT	CCT	TCA		840	
																					280	
ACG	GTT	TTC	GCT	GCC	TAC	AAT	TAC	TAC	GCT	GGA	CCG	AAG	GAA	ATC	AGA	ATC	TAT	CCG	TAC		900	
																					300	
AAC	AAC	CAC	GAG	GGA	GGA	GGA	GCG	TCT	TTC	CAA	GCG	GTT	GAA	CAG	GTG	AAA	TTC	TTG	AAA		960	
																					320	
CTA	TTT	GAG	AAA	GGC	TAA																978	
																					326	
Leu	Phe	Glu	Lys	Gly																		

Figure 20

Polyangium brachysporum Esterase 78mc1

TTG AAG TAC TTC AAA GCC CGG CTT GCC GGC ATC ACC TTG CTC GGC CTG CTG GCC TGC ACC	60
Leu Lys Tyr Phe Lys Ala Arg Leu Ala Gly Ile Thr Leu Leu Gly Leu Leu Ala Cys Thr	20
TCG GCC TCG GCG CAG ACC GAG CCC ATC GTG TTC GTG CAC GGC TAT TCC GGC AGC GCA TCC	120
Ser Ala Ser Ala Gln Thr Glu Pro Ile Val Phe Val His Gly Tyr Ser Gly Ser Ala Ser	40
AAC TGG GAC ACC ATG CTG GGC CGC TTC CGG TCG AAC GGT TAT GCG TCC GGC TCG CTC TAC	180
Asn Trp Asp Thr Met Leu Gly Arg Phe Arg Ser Asn Gly Tyr Ala Ser Gly Ser Leu Tyr	60
ACC TTC AAC TAC AAC TCG TTG GTC AGC AGC AAC CGC ACC AGC GCC AGC GAG CTG CGC AGC	240
Thr Phe Asn Tyr Asn Ser Leu Val Ser Ser Asn Arg Thr Ser Ala Ser Glu Leu Arg Ser	80
TTC GTC AAC ACC GTG CGT TCG CGC CAC GGC AAC GCC CGC ATC GCG CTG GTC GCC CAC TCC	300
Phe Val Asn Thr Val Arg Ser Arg His Gly Asn Ala Arg Ile Ala Leu Val Ala His Ser	100
AAC GGC GGG CTG GTG TCG CGC TGG TAT CGC GCG GAG CTG GGC GGC GAA ACG GCC ACC CGC	360
Asn Gly Gly Leu Val Ser Arg Trp Tyr Arg Ala Glu Leu Gly Glu Thr Ala Thr Arg	120
CGC TTC GTG ACG CTG GGC ACG CGC CAC CGG GGC ACC ACC TGG GCC TAT GCG TGC TAC AGC	420
Arg Phe Val Thr Leu Gly Thr Pro His Arg Gly Thr Thr Trp Ala Tyr Ala Cys Tyr Ser	160
CCC GCA TGT TTC GAG ATG CGC CCC GGC TCC AGC TTG CTG ACC ACG CTG GGC TCG CGT GCC	480
Pro Ala Cys Phe Glu Met Arg Pro Gly Ser Ser Leu Leu Thr Thr Leu Gly Ser Arg Ala	160
TGC GAC CGC TCG CTG TGG TCG AAC ACC GAC GGC ATC ATC CTG CGG GCG TCC AGC GCG CGG	540
Cys Asp Arg Ser Leu Trp Ser Asn Thr Asp Gly Ile Ile Leu Pro Ala Ser Ser Ala Gln	180
TGT GGT GTC AGC ACG CGC ACT GGC GAC GTC AGC CAT CTC GAC CTG CTG ACC GAC TCT CGC	600
Cys Gly Val Ser Thr Arg Thr Ala Asp Val Ser His Leu Asp Leu Leu Thr Asp Ser Arg	200
GTC TAC ACG CAG TTG CGC ACG CAG TTG CAA TGA CGG TGA CGG TGC ACC GAA CGT GCA CCT	660
Val Tyr Thr Gln Leu Arg Thr Gln End Gly End Arg Cys Thr Glu Arg Ala Pro	220
	661

G